```
OM of: US-09-652-292-2 to: N_Geneseq_1101:*   out_format : pfs

Date: Mar 15, 2002  7:56 AM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+_p2n.model -DEV=xlh
 -Q=/cgn2_1/USPTO_spool/US09652292/runat_13032002_161726_3249/app_query.fasta_1.606
 -DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=rng -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cd1 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09652292_@CGN1_1_238 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-652-292-2
Query length: 541
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 111.030000



score_list:
Sequence                Strd Orig    ZScore    EScore Len  ! Documentation  ..
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX35516 +  567.00  670.76  2.8e-29   2471  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55868 +  556.00  658.66  1.3e-28   2177  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55869 +  547.00  646.52  6.2e-28   2504  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC42332 +  541.00  642.19  1.1e-27   1826  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH88793 +  517.50  613.48  4.3e-26   1905  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ24474 +  507.00  601.03  2.1e-25   1862  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC36954 +  506.00  598.57  2.9e-25   2134  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA65400 +  504.00  599.55  2.6e-25   1476  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH65168 +  502.00  597.16  3.5e-25   1473  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH64966 -  502.00  546.88  2.2e-22  349980 !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32202 +  500.50  592.72  6.1e-25   1960  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC50883 +  488.50  578.86  3.6e-24   1835  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55866 +  486.00  574.67  6.2e-24   2087  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC37492 +  482.50  571.62  9.2e-24   1837  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD09552 +  478.00  565.05  2.1e-23   2080  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55865 +  470.00  554.82  7.9e-23   2217  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC43261 +  466.50  554.03  8.8e-23   1527  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:AAV11474 +  457.50  535.91  9.0e-22   3370  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55867 +  456.00  538.57  6.4e-22   2072  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55870 +  455.00  540.09  5.2e-22   1541  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32200 +  438.50  518.89  7.9e-21   1776  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC50050 +  433.50  514.89  1.3e-20   1425  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT66496 +  418.50  493.96  1.9e-19   1943  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32203 +  417.00  493.10  2.2e-19   1752  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH68347 +  414.00  490.76  2.9e-19   1524  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH68534 -  414.00  441.93  1.5e-16  309400 !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55871 +  407.50  480.38  1.1e-18   2011  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC33376 +  406.50  481.83  9.2e-19   1507  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32195 +  404.50  478.44  1.4e-18   1675  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ93709 +  397.00  446.44  8.6e-17  20383  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC39099 +  392.00  463.33  9.9e-18   1683  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC36319 +  390.00  460.61  1.4e-17   1741  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC46009 +  387.50  458.13  1.9e-17   1642  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC40459 +  384.50  453.91  3.3e-17   1755  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC45857 +  382.00  453.00  3.7e-17   1395  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF81396 +  381.50  446.71  8.3e-17   2592  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC50049 +  379.50  449.79  5.6e-17   1425  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF14638 +  379.00  446.57  8.5e-17   1895  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF98714 +  376.50  439.67  2.1e-16   2893  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:AAT66495 +  375.50  442.75  1.4e-16   1815  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC65871 +  375.50  438.58  2.4e-16   2856  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC45298 +  373.50  441.02  1.7e-16   1685  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC45290 +  373.50  440.71  1.8e-16   1743  !
/SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:AAQ11148 +  371.50  434.67  3.9e-16   2587  !
/SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH53274 +  356.00  424.84  1.4e-15    987  !

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAX35516

seq_documentation_block:
ID   AAX35516 standard; cDNA; 2471 BP.
XX
AC   AAX35516;
XX
DT   08-JUL-1999  (first entry)
XX
DE   cDNA encoding facilitative glucose transporter protein GLUT8.
XX
KW   Facilitative glucose transporter protein; GLUT8; malignancy;
KW   breast cancer; prostate cancer; epithelial cell cancer;
KW   non-insulin-dependent diabetes mellitus; insulin resistance;
KW   central obesity; hypertension; dyslypidaemia; glucose intolerance;
KW   cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9918125-A1.
XX
PD   15-APR-1999.
XX
PF   30-SEP-1998;   98WO-AU00819.
XX
PR   01-OCT-1997;   97AU-0009573.
XX
PA   (SVIN-) ST VINCENTS INST MEDICAL RES.
XX
PI   Best JD,  Rogers SD;
XX
DR   WPI; 1999-277253/23.
DR   P-PSDB; AAY02168.
XX
PT   Nucleic acid encoding a facilitative glucose transporter
XX
PS   Claim 4; Page 61-62; 72pp; English.
XX
CC   The present sequence encodes a facilitative glucose transporter
CC   protein, GLUT8. Agents that inhibit activity or expression of
CC   GLUT8 (particularly non-utilizable glucose analogues, antisense
CC   sequences or dominant negative mutants) are used to treat
CC   malignancy, particularly cancer of breast, prostate and epithelial
CC   cells (e.g. skin or colon). Agents that upregulate expression of GLUT8
CC   (e.g. the GLUT8 gene, administered by tissue-localized gene therapy) are
CC   used to treat non-insulin-dependent diabetes mellitus and/or insulin
CC   resistance (e.g. central obesity, hypertension, dyslypidaemia or glucose
CC   intolerance). Detecting expression of GLUT8 is used for diagnosis,
CC   monitoring and staging of cancers, particularly of the breast. Antibodies
CC   raised against GLUT8 are useful as immunoassay reagents and as
CC   therapeutic inhibitors.
XX
SQ   Sequence 2471 BP; 611 A; 613 C; 526 G; 678 T; 43 other;


alignment_scores:
            Quality:  567.00              Length:    563
              Ratio:  1.989                 Gaps:     19
 Percent Similarity:  50.622   Percent Identity: 31.261

alignment_block:
 US-09-652-292-2 x AAX35516   ..

 Align seg 1/1  to: AAX35516  from: 1  to: 2471

       6 ProValLeuProLeuCysAlaSerValSerLeuLeuGly...GlyLeuTh 21
         |||::::::   :::   :::|||:::   |||||||||   |||||::
```

```
 229 CCAGCATTTAGGATTTCTTCCTCTATAATCTTGCTGGGTGCTGGTCTTGC 278

  21 r.....PheGlyTyrGluLeuAlaValIleS 30
     :                 ||||||   ::::::   |||   :
 279 AGGGCCATCTACTGGGGATAGGTGGTTTGGGGTCTCAGTG...GTGGGCA 325

  30 erGly.....AlaLeuLeuProLeuGlnLeuAsp 39
     ::|||              |||||||||   :::|||
 326 CCGGCTTGTTCTTGCCTCCTCTGCAGCTCCTCTTGCCGCCTCGCCTGCTG 375

  40 PheGlyLeuSerCysLeuGluGlnGluPhe.....LeuValGl 52
     |||      ::: . ||||||:::                |||
 376 TTCACTCATGCAATCCTTGAACGCCTGCACCTGTGGCTGGCATTGCCGCC 425

  52 ySerLeuLeuLeuGly...AlaLeuLeuAlaSerLeuValGlyGlyPheL 68
       |||:::||||||   |||||||||     ||||||||||
 426 AGTCCTGGTGCTGGGCCATGCACTCCTGCACTGCAAAGTGGGAGGC.... 471

  68 euIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuVal 84
                     :::   |||      |||::::::||||||
 472 .....AGCACAGCCAGAGCGGGAGATCAGCTGGTC 501

  85 LeuLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLe 101
        :::      |||   :::   |||                |||::
 502 CAGCGGGTCCTCCTCCTCATCGTCTTTCTTCAC.....CGTTGGGT 542

 101 u.....ValLeuGlyArgAlaValV 108
     :                :::||||||
 543 CCAGGTATGGCCTTGAGGGACTGAGGTTGACATCCTGGGG..... 582

 108 alGlyPheAlaIleSerLeuSerSerMetAlaCysCysIleTyrValSer 124

 582 ..................................................... 582

 125 GluLeuValGlyProArgGlnArgGlyValLeuValSerLeuTyrGluAl 141
        ::|||   |||      |||                    :::
 583 ......ATGGGGAGTCGAACAGGTGGGAGAAGAGGGCCGGAACTCCGGCC 626

 141 aGlyIleThrValGlyIleLeuLeuSerTyrAlaLeuAsnTyrAlaLeuA 158
      |||:::   ::::::||||||   :::|||      ||||||||||:::|
 627 GGGATTCCGGATCAGCATTCTTTCTGCCTATATTTCAAATTACGCATTTG 676

 158 laGlyThrProTrpGlyTrpArgHisMetPheGlyTrpAlaThrAlaPro 174
     ||::::::      ||||||::::::|||||||||   :::
 677 CCAATGTTTTCCATGGCTGGAAGTACATGTTTGGTCTTGTGATTCCCTTG 726

 175 AlaValLeuGlnSerLeuSerLeuLeuPheLeuProAlaGlyThrAspGl 191
     :::||||||||:::::::::::::   |||||||||   :::
 727 GGAGTTTTGCAAGCAATTGCAATGTATTTTCTTCCTCCAAGC........ 768

 191 uThrAlaThrHisLysAspLeuIleProLeuGlnGly......GlyGluA 206
                     ::::::   :::::::|||      |||   |
 769 .............CCTCGGTTTCTGGTGATGAAAGGACAAGAGGGAGCTG 805

 206 laProLysLeuGlyProGlyArgProArg..................... 215
     ||   |||:::      ||||||   |||
 806 CTAGCAAGGTT...CTTGGAAGGTTAAGAGCACTCTCAGATACAACTGAG 852

 216 .....................................TyrSerPh 218
                                          ||||||||
 853 GAACTCACTGTGATCAAATCCTCCCTGAAAGATGAATATCAGTACAGTTT 902

 218 eLeuAspLeuPheArgAlaArgAspAsnMetArgGlyArgThrThrValG 235
     |   ||||||||||||:::::::||||||||||||   |||      :::|
 903 TTGGGATCTGTTTCGTTCAAAAGACAACATGCGGACCCGAATAATGATAG 952

 235 lyLeuGlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnValLeu 251
     |||||   ||||||:::|||   |||:::||||||||||||||||:::|||
 953 GACTAACACTAGTATTTTTTGTACAAATCACTGGCCAACCAAACATATTG 1002

 252 CysTyrAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerSe 268
     ||||||||||||:::::::::||||||||||||:::::::::::::::
1003 TTCTATGCATCAACTGTTTTGAAGTCAGTTGGATTTCAAAGCAATGAGGC 1052

 268 rAlaValLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThrL 285
     :|||   |||||||||:::|||:::|||:::|||||||||   :::||||:
1053 AGCTAGCCTCGCCTCCACTGGGGTTGGAGTCGTCAAGGTCATTAGCACCA 1102

 285 euThrAlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLeu 301
     ::   |||      ||||||||||::::::|||   ::::::::::|||
1103 TACCTGCCACTCTTCTTGTAGACCATGTCGGCAGCAAAACATTCCTCTGC 1152

 302 AlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValSe 318
       |||                |||      :::|||::::::   :::::
1153 ATTGGC.............TTGCTAAATGCTGGATTAAGCCACACTGA 1187

 318 rPheAlaValProMetAspSerGlyProSerCysLeuAlaValProAsnA 335
     ::::   :::      |||
1188 ATACCAGATAGTCACAGAC............................. 1206

 335 laThrGlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSer 351
                        |||||||||
1207 ...................CCTGGGGAC..................... 1215

 352 SerLeuProProIleProArgThrAsnGluAspGlnArgGluProIleLe 368
                 :::|||
1216 ............GTCCCA............................... 1221

 368 uSerThrAlaLysLysThrLysProHisProArgSerGlyAspProSerA 385

1221 ..................................................... 1221

 385 laProProArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuPro 401

1221 ..................................................... 1221

 402 AlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetVa 418
                 |||:::|||:::|||   :::|||      |||:::||
1222 ............GCTTTTTTGAAATGGCTGTCCTTAGCCAGCTTGCTTGT 1259

 418 lPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValL 435
     |:::|||:::||||||||||:::|||:::||||||:::   ||||||||||
1260 TTATGTTGCTGCTTTTTCAATTGGTCTAGGACCAATGCCCTGGCTGGTGC 1309

 435 euSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCys 451
     ||||||||||||:::|||      ||||||||||||||||:::|||:::
1310 TCAGCGAGATCTTTCCTGGTGGGATCAGAGGACGAGCCATGGCTTTAACT 1359

 452 AsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAs 468
     :::|||:::||||||:::   ||||||:::||||||||:::||||||
1360 TCTAGCATGAACTGGGGCATCAATCTCCTCATCTCGCTGACATTTTTG.. 1407

 468 pLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeuT 485
       ::::::   |||||||||   |||:::   :::::::|||   :::
1408 .ACTGTAAATCTTATTGGCCTGCCATGGGTGTGCTTTATATATACAATCA 1456

 485 hrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLys 501
       ::::::   :::|||   |||:::   :::|||:::||||||||||||
1457 TGAGTCTAGCATCCCTGCTTTTTGTTGTTATGTTTATACCTGAGACAAAG 1506

 502 GlyGlnSerLeuAlaGluIleAspGlnGlnPheGlnLys 514
     |||   ||||||   :::|||:::::::::::   |||
1507 GGATGCTCTTTGGAACAAATATCAATGGAGCTAGCAAAA 1545
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55868

seq_documentation_block:
ID AAF55868 standard; cDNA; 2177 BP.
XX
AC AAF55868;
XX
DT 17-APR-2001 (first entry)
XX
DE Human GLUTX2 coding sequence.
XX
KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;

```
KW   hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW   hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS   Homo sapiens.
XX
PN   WO200104145-A2.
XX
PD   18-JAN-2001.
XX
PF   14-JUL-2000; 2000WO-IB01042.
XX
PR   14-JUL-1999;   99US-0143907.
PR   27-AUG-1999;   99US-0151140.
PR   23-FEB-2000;  2000US-0184285.
PR   13-JUL-2000;  2000US-0616132.
XX
PA   (UYLA-) UNIV LAUSANNE.
XX
PI   Thorens B,  Ibberson M,  Uldry M;
XX
DR   WPI; 2001-112615/12.
DR   P-PSDB; AAB66935.
XX
PT   Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT   the prevention, diagnosis and treatment of hexose transport disorders,
PT   e.g. ischemia and diabetes -
XX
PS   Claim 3; Page 76-77; 124pp; English.
XX
CC   The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC   AAB66932-AAB66941). The GLUTX proteins are related to the facultative
CC   glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC   function. The GLUTX proteins may be used in the diagnosis, prevention and
CC   treatment of hexose transport disorders such as ischaemia, diabetes,
CC   hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC   neurodegenerative disease. The present sequence is the coding sequence
CC   for human GLUTX2.
XX
SQ   Sequence 2177 BP; 473 A; 568 C; 618 G; 518 T; 0 other;


alignment_scores:
            Quality:  556.00              Length:    596
              Ratio:  1.680                Gaps:     15
 Percent Similarity:  55.537   Percent Identity: 30.537

alignment_block:
 US-09-652-292-2 x AAF55868   ..

 Align seg 1/1  to: AAF55868  from: 1  to: 2177

       4 SerProProValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLe 20
         :::|||      :::   :::   |||      |||   |||||||||::
     459 ACCCCCGCCTTCGTGTACGTGGTGGCCGTCTTCTCCGCGCTGGGCGGCTT 508

      20 uThrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeuLeuProLeuG 37
         :   ||||||||||:::   :::|||:::||||||||||:::|||   |||:
     509 CCTGTTTGGCTATGACACCGGGGTGGTGTCAGGGGCCATGCTGCTGCTCA 558

      37 lnLeuAspPheGlyLeuSerCysLeuGluGlnGluPheLeuValGlySer 53
         ::   ::::::::::|||:::::|||   |||||||:::|||||||:::|||
     559 AGCGGCAGCTCAGTCTGGACGCGCTGTGGCAGGAGCTGCTGGTGTCCAGC 608

      54 LeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLeuIleAs 70
            :::   :::|||   :::::::::|||:::||||||   |||
     609 ACGGTGGGGGCGGCTGCCGTCTCGGCGCTGGCCGGAGGCGCCCTCAACGG 658

      70 pCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeuA 87
            :::|||||||:::   ||||||||||   ::::::   ::::::   |
     659 CGTCTTCGGCCGCCGCGCTGCCATCCTCCTGGCCAGTGCCCTCTTCACCG 708

      87 laGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuValLeu 103
         ||||||||   :::|||:::   |||::::::         |||:::

     709 CCGGCTCCGCGGTGCTGGCTGCGGCCAACAACAAGGAGACACTGCTCGCC 758

     104 GlyArgAlaValValGlyPheAlaIleSerLeuSerSerMetAlaCysCy 120
         ||||||   ||||||||||::::::|||:::::::::|||||||:::
     759 GGCCGCCTGGTCGTGGGACTCGGCATCGGCATTGCTTCTATGACAGTGCC 808

     120 sIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuValS 137
          :::|||::::::|||:::      |||:::   ||||||   ||||||:
     809 AGTGTACATTGCGGAGGTCTCACCACCCAATTTAAGAGGCCGATTAGTCA 858

     137 erLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyrAlaLeu 153
         :::::            ||||||   |||   :::::::::   ::::::
     859 CCATTAATACCCTCTTCATCACAGGAGGGCAGTTCTTTGCAAGTGTTGTT 908

     154 AsnTyrAlaLeuAlaGlyThrProTrp...GlyTrpArgHisMetPheGl 169
         :::   |||::::::               |||||||||:::|||:::||
     909 GATGGAGCCTTCAGTTATCTCCAGAAGGATGGATGGAGGTACATGTTGGG 958

     169 yTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPheLeuP 186
         |   |||::::::|||||||||:::|||   :::::::::|||||||||
     959 ACTTGCAGCAGTTCCGGCGGTTATACAGTTTTTTGGCTTTCTCTTTTTGC 1008

     186 roAlaGly..................................... 188
         ||   :::
    1009 CTGAAAGCCCTCGATGGCTTATTCAGAAAGGACAGACTCAGAAGGCCCGT 1058

     189 .................................ThrAspGluThrAl 193
                                                 :::|||
    1059 AGAATTTTATCTCAGATGCGTGGTAACCAGACCATTGATGAGGAATATGA 1108

     193 aThrHisLysAspLeuIleProLeuGlnGlyGlyGluAlaProLysLeuG 210
          :::   |||:::   |||    :::     |||:::   :::   |
    1109 TAGCATCAAAAACAACATTGAAGAGGAGGAAAAAGAGGTTGGCTCAGCTG 1158

     210 lyProGlyArgProArg...TyrSerPheLeuAspLeuPheArgAlaArg 225
         |||||         |||      |||:::            ||||||
    1159 GACCTGTGATCTGCAGAATGCTGAGTTATCCCCCAACTCGCCGAGCT... 1205

     226 AspAsnMetArgGlyArgThrThrValGlyLeuGlyLeuValLeuPheGl 242
                                 ||||||   ||||||   :::|||||
    1206 .................TTAATTGTGGGTTGTGGCCTACAAATGTTCCA 1237

     242 nGlnLeuThrGlyGlnProAsnValLeuCysTyrAlaSerThrIlePheS 259
         |||||||:::|||      :::::::::  |||::::::||||||::::
    1238 GCAGCTCTCAGGCATTAACACCATCATGTACTACAGTGCAACCATTCTGC 1287

     259 erSerValGlyPheHisGlyGlySerSerAlaVal...LeuAlaSerVal 274
         ::      |||   :::           |||:::   ||||||||||||
    1288 AGATGTCTGGTGTTGAAGATGATAGACTTGCAATATGGCTGGCTTCAGTT 1337

     275 GlyLeuGlyAlaValLysValAlaAlaThrLeuThrAlaMetGlyLeuVa 291
               :::   ::::::         ||||||:::::::   |||||
    1338 ...ACAGCCTTCACAAATTTCATTTTCACACTTGTGGGAGTCTGGCTTGT 1384

     291 lAspArgAlaGlyArgArgAlaLeuLeu.........LeuAlaGlyCysA 305
         |::::::::::|||||||||   |||           |||||||||   :
    1385 TGAGAAGGTGGGCCGCAGAAAGCTTACCTTTGGTAGTTTAGCAGGT...A 1431

     305 laLeuMetAlaLeuSerValSerGlyIleGlyLeuValSerPheAla... 320
         ::   :::||||||   :::   ::::::|||:::|||      |||
    1432 CCACCGTAGCACTCATTATTCTTGCCTTGGGATTTGTGCTATCAGCCCAA 1481

     321 .............................ValProMetAspSerGlyPr 327
                                         :::|||      ::::::
    1482 GTTTCCCCACGCATCACTTTTAAGCCAATAGCTCCGTCAGGTCAGAACGC 1531

     327 oSerCysLeuAlaValProAsnAlaThrGlyGlnThrGlyLeuProGlyA 344
          :::|||            :::      ::::::      |||      |
    1532 CACTTGCACAAGATAC...AGTTACTGTAATGAATGTATGTTGGATCCAG 1578

     344 spSerGlyLeu.....................LeuGlnAspSerSer 352
         ||   |||:::                        :::   |||||||||
    1579 ACTGCGGTTTCTGCTACAAGATGAACAAATCAACTGTCATTGACTCCTCC 1628
```

```
 353 LeuProProIleProArg......ThrAsnGluAsp..... 362
1629 TGTGTTCCAGTTAATAAAGCATCTACAAATGAGGCAGCCTGGGGCAGGTG 1678

 363 .GlnArgGluProIleLeuSerThr.....AlaLysLysT 374
1679 TGAAAATGAAACCAAGTTCAAAACAGAAGATATATTTTGGGCTTACAATT 1728

 374 hrLysProHisProArgSerGlyAspProSerAlaProProArgLeuAla 390
1729 TCTGCCCTACTCCATACTCC..... 1748

 391 LeuSerSerAlaLeuProGlyProProLeuProAlaArgGlyHisAlaLe 407

1748 ..... 1748

 407 uLeuArgTrpThrAlaLeuLeuCysLeuMetValPheValSerAlaPheS 424
1749 .......TGGACTGCACTTCTGGGCCTTATTTTATATCTTGTCTTCTTTG 1791

 424 erPheGlyPheGlyProValThrTrpLeuValLeuSerGluIleTyrPro 440
1792 CACCTGGAATGGGACCAATGCCTTGGACTGTGAATTCTGAAATATATCCC 1841

 441 ValGluIleArgGlyArgAlaPheAlaPheCysAsnSerPheAsnTrpAl 457
1842 CTTTGGGCAAGAAGTACAGGAAATGCATGTTCATCTGGAATAAACTGGAT 1891

 457 aAlaAsnLeuPheIleSerLeuSerPheLeuAspLeuIleGlyThrIleG 474
1892 TTTCAATGTCCTGGTTTCACTAACATTTTTACACACAGCAGAGTATCTTA 1941

 474 lyLeuSerTrpThrPheLeuLeuTyrGlyLeuThrAlaValLeuGlyLeu 490
1942 CATACTATGGAGCTTTCTTCCTCTATGCTGGATTTGCTGCTGTGGGACTC 1991

 491 GlyPheIleTyrLeuPheValProGluThrLysGlyGlnSerLeuAlaGl 507
1992 CTTTTCATCTATGGCTGTCTTCCTGAGACCAAAGGCAAAAAATTAGAGGA 2041

 507 uIleAspGlnGlnPheGlnLysArgArgPheThrLeuSerPheGlyHisA 524
2042 AATTGAATCACTCTTTGACAACAGGCTATGTACATGTGGCACTTCAGATT 2091

 524 rgGlnAsnSerThrGlyIleProTyrSerArgIleGlu 536
2092 CTGATGAAGGGAGATATATTGAATATATCCGCGTGAAG 2129
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55869

seq_documentation_block:

ID AAF55869 standard; cDNA; 2504 BP.
XX
AC AAF55869;
XX
DT 17-APR-2001 (first entry)
XX
DE Rat GLUTX2 coding sequence.
XX
KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Rattus sp.
XX
PN WO200104145-A2.
XX
PD 18-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-IB01042.
XX
PR 14-JUL-1999; 99US-0143907.
PR 27-AUG-1999; 99US-0151140.
PR 23-FEB-2000; 2000US-0184285.
PR 13-JUL-2000; 2000US-0616132.
XX
PA (UYLA-) UNIV LAUSANNE.
XX
PI Thorens B, Ibberson M, Uldry M;
XX
DR WPI; 2001-112615/12.
DR P-PSDB; AAB66936.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT the prevention, diagnosis and treatment of hexose transport disorders,
PT e.g. ischemia and diabetes -
XX
PS Claim 3; Page 77-79; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is the coding sequence
CC for rat GLUTX2.
XX
SQ Sequence 2504 BP; 547 A; 713 C; 693 G; 551 T; 0 other;

alignment_scores:

Quality:	547.00	Length:	585
Ratio:	1.609	Gaps:	15
Percent Similarity:	58.120	Percent Identity:	30.256

alignment_block:
US-09-652-292-2 x AAF55869 ..

Align seg 1/1 to: AAF55869 from: 1 to: 2504

```
   4 SerProProValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLe 20
 283 ACCCCCGCCTTCGTGTACGCGGCCGCCGCCTTCTCCGCTCTCGGCGGCTT 332

  20 uThrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeuLeuProLeuG 37
 333 CCTGTTCGGCTACGACACCGGCGTGGTGTCGGGGGCCATGCTGCTGCTGC 382

  37 lnLeuAspPheGlyLeuSerCysLeuGluGlnGluPheLeuValGlySer 53
 383 GGCGCCAGATGCGCCTGGGCGCGATGTGGCAGGAGCTGCTGGTGTCGGGC 432

  54 LeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLeuIleAs 70
 433 GCGGTGGGCGCGGCCGCCGTCGCGGCGCTGGCCGGAGGGGCCCTGAACGG 482

  70 pCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeuA 87
 483 CGCCCTCGGTCGGCGAAGCGCCATCCTGCTGGCCAGCGCCCTGTGCACCG 532

  87 laGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuValLeu 103
 533 TGGGCTCCGCCGTGCTGGCCGCCGCCGCCAACAAGGAGACGCTGCTGGCC 582

 104 GlyArgAlaValValGlyPheAlaIleSerLeuSerSerMetAlaCysCy 120
 583 GGCCGCCTGGTCGTGGGGCTCGGCATCGGCATCGCATCCATGACAGTGCC 632

 120 sIleTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuValS 137
 633 CGTGTACATCGCCGAGGTCTCTCCACCCAACCTGAGAGGTCGCTTGGTCA 682

 137 erLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyrAlaLeu 153
```

```
 683 CCATCAACACCCTCTTCATCACCGGCGGACAGTTCTTTGCGAGCGTTGTT 732

 154 AsnTyrAlaLeuAlaGlyThrProTrp...GlyTrpArgHisMetPheGl 169
     :::   |||::::::               |||||||||:::|||:::||
 733 GATGGAGCCTTTAGTTACCTGCAGAAGGATGGATGGAGGTACATGTTGGG 782

 169 yTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPheLeuP 186
     |   |||:::   |||||||||:::|||   |||::::::||||||||||
 783 ACTTGCGGCCATTCCAGCCGTTATACAATTCCTCGGATTCCTCTTTTTGC 832

 186 roAlaGly..................................... 188
     ||   :::
 833 CCGAAAGTCCTCGGTGGCTGATACAGAAAGGACAGACTCAGAAGGCCCGC 882

 189 .................................ThrAspGluThrAl 193
                                        :::|||
 883 CGAATTTTGTCCCAGATGCGTGGGAATCAGACCATTGACGAGGAGTATGA 932

 193 aThrHisLysAspLeuIleProLeuGlnGlyGlyGluAlaProLysLeuG 210
       :::   ::::::   |||      :::      ||||||          |
 933 CAGCATCAGGAACAGCATCGAGGAGGAGGAGAAGGAAGCCAGCGCAGCTG 982

 210 lyProGlyArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAsp 226
     |||||                  :::::::      :::   :::
 983 GACCT.................ATAATCTGCAGAATGCTGAGTTACCCC 1014

 227 AsnMetArgGlyArgThrThrValGlyLeuGlyLeuValLeuPheGlnGl 243
         |||          :::||||||   ||||||   ::::||||||||
1015 CCAACTCGCCGAGCGTTAGCTGTGGGATGTGGCTTACAGATGTTCCAGCA 1064

 243 nLeuThrGlyGlnProAsnValLeuCysTyrAlaSerThrIlePheSerS 260
     ||||:::|||      :::::::::   |||::::::||||||::::::
1065 GCTCTCGGGCATCAACACTATCATGTACTACAGCGCGACCATCCTGCAGA 1114

 260 erValGlyPheHisGlyGlySerSerAlaVal...LeuAlaSerValGly 275
         |||   :::          |||:::   ||||||||||:::
1115 TGTCCGGCGTGGAAGATGATAGACTTGCAATATGGCTGGCTTCCATT... 1161

 276 LeuGlyAlaValLysValAlaAlaThrLeuThrAlaMetGlyLeuValAs 292
       :::   ::::::         ||||||:::::::::   |||||||::
1162 ACAGCGTTCACCAATTTCATTTTCACACTGGTGGGCGTCTGGCTCGTGGA 1211

 292 pArgAlaGlyArgArgAlaLeuLeu.........LeuAlaGlyCysAlaL 306
     :::::::|||||||||   |||            |||||||||   :::
1212 GAAGGTGGGCCGCAGGAAGCTCACCTTTGGCAGTTTGGCAGGT...ACCA 1258

 306 euMetAlaLeuSerValSerGlyIleGlyLeu.................. 316
       :::|||||||::::::   :::::::|||:::
1259 CAGTAGCACTTACAATTCTTGCCTTGGGATTTCTGCTCTCAGCTCAGGTC 1308

 317 .........ValSerPheAlaValProMetAspSerGlyProSerCysLe 330
             |||:::|||               ||||||   ::::::
1309 TCACCACGGGTCACTTTCAGACCAACGGCTCCCTCGGGTCAAAATGCCAC 1358

 330 u.AlaValProAsnAlaThrGly.................GlnThrGly 340
       |||      ::::::|||                    |||||||:::
1359 CTGCACAGAATACAGTTACTGTAATGAGTGTATGCTGGATCCAGACTGCG 1408

 341 LeuProGlyAspSerGlyLeuLeuGlnAspSerSerLeuProProIlePr 357
     :::   :::      :::   ::::::   |||   :::|||||||:::
1409 GTTTCTGCTACAAGATCAACAGTTCGGCTGTCATTGATTCCTCCTGTGTT 1458

 357 oArgThrAsnGlu......AspGlnArgGluProIleLeuSerThrAlaL 372
       |||   ::::::         ::::::::::|||      :::::::::
1459 CCGGTGAACAAAGCTTCCACCAATGAAGCAGCCT......GGGGCAGGTG 1502

 372 ysLysThrLysProHisProArgSerGlyAspProSerAlaProProArg 388
       |||||||||                  ||||||:::      :::
1503 TGAAAACGAAA.....................CCAAGTTCAAAGCAGAAG 1531

 389 LeuAlaLeuSer.........SerAlaLeuProGlyProProLeuProAl 402
     :::      :::         ||||||||||||
1532 ATGTTCACTGGGCTTACAGTTTCTGCCCTACCCCA.............. 1566

 402 aArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetValP 419
                              ||||||||||||:::   |||:::::::
1567 ..............TACTCCTGGACAGCACTCGTGGGCCTGGTGTTAT 1600

 419 heValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValLeu 435
     :::::      |||:::   |||:::||||||:::   |||   |||
1601 ATCTTGTTTTCTTTGCACCCGGAATGGGACCAATGCCCTGGACCGTGAAC 1650

 436 SerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCysAs 452
     ||||||||||||||:::      |||:::   :::   |||
1651 TCTGAAATCTATCCTCTCTGGGCAAGAAGTACAGGAAATGCGTGCTCAGC 1700

 452 nSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAspL 469
      ::::::||||||      |||:::::::::||||||:::||||||
1701 TGGAATAAACTGGATTTTCAACGTGCTGGTTTCACTGACCTTTTTACACA 1750

 469 euIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeuThr 485
                :::            :::|||:::||||||:::
1751 CAGCGGAGTATCTTACATACTATGGAGCGTTCTTCCTCTACGCCGGATTC 1800

 486 AlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLysGl 502
     |||::::::||||||   |||:::|||      :::||||||||||||||
1801 GCCGCTGTGGGACTGCTTTTCGTCTATGGCTGTCTTCCTGAAACCAAAGG 1850

 502 yGlnSerLeuAlaGluIleAspGlnGlnPheGlnLysArgArgPheThrL 519
     |:::::::|||   ||||||::::::   |||:::   |||      |||
1851 GAAAAAACTAGAGGAAATCGAATCGCTCTTCGACCACCGGCTGTGCACCT 1900

 519 euSerPheGlyHisArgGlnAsnSerThrGlyIleProTyrSerArgIle 535
       :::   :::      :::::::::      |||   |||   |||:::
1901 GCGGCACCGCGGACTCGGACGAGGGCAGGTACATCGAGTACATCCGCGTG 1950

 536 Glu 536
     :::
1951 AAG 1953
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC42332

seq_documentation_block:

ID AAC42332 standard; DNA; 1826 BP.
XX
AC AAC42332;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35149.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

```
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
            Quality:  541.00              Length:    559
              Ratio:  1.639                Gaps:     13
 Percent Similarity:  59.034   Percent Identity: 28.801

alignment_block:
 US-09-652-292-2 x AAC42332   ..

 Align seg 1/1  to: AAC42332  from: 1  to: 1826

      4 SerProProValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLe 20
        :::|||   ::::::   |||:::   |||::::::   :::||||||||
    147 ACACCTTACATCATGCGACTTGCTCTCTCCGCCGGA...ATCGGAGGTCT 193

     20 uThrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeuLeuProLeuG 37
        |   ||||||||||:::   :::||||||||||||||||||||||   ::::
    194 TCTCTTCGGTTACGATACCGGAGTCATTTCCGGAGCTCTTCTTTTCATCA 243

     37 lnLeuAspPhe.........GlyLeuSerCysLeuGluGlnGluPheLeu 50
        ::   ||||||          :::          |||:::   :::
    244 AAGAAGATTTTGATGAAGTTGATAAGAAAACATGGCTTCAGTCAACTATT 293

     51 ValGlySerLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPh 67
        |||:::      :::   ||||||:::::::::::   |||||||||::
    294 GTTAGTATGGCAGTGGCTGGAGCCATCGTCGGAGCAGCAGTCGGAGGTTG 343

     67 eLeuIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuV 84
        ::::   |||   :::||||||:::::::::||||||   ::::::::::
    344 GATCAATGATAAGTTTGGTAGGAGGATGTCGATTCTTATCGCCGATGTTC 393

     84 alLeuLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrp 100
        ::::::|||   |||:::::::::::::::::::|||   :::
    394 TTTTCTTGATCGGTGCGATTGTGATGGCATTTGCTCCGGCTCCTTGGGTT 443

    101 LeuValLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerMe 117
        :::::::::||||||      |||||||||:::::::::::::::||||||
    444 ATCATCGTTGGAAGAATCTTTGTTGGGTTTGGTGTTGGTATGGCTTCGAT 493

    117 tAlaCysCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyV 134
        |:::      :::|||:::||||||          |||   ||||||:
    494 GACGTCTCCGTTGTATATCTCTGAAGCTTCTCCGGCGAGGATTAGAGGTG 543

    134 alLeuValSerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSer 150
        ::|||||||||               ||||||   |||   ::::::|||
    544 CACTTGTTAGTACCAATGGTCTTCTCATCACTGGTGGACAATTCTTCTCG 593

    151 TyrAlaLeuAsnTyrAlaLeuAlaGlyThrProTrpGlyTrpArgHisMe 167
        |||   :::|||   |||::::::   ||||||      ||||||   ||
    594 TACCTCATCAATCTTGCTTTTGTCCATACTCCGGGAACATGGAGATGGAT 643

    167 tPheGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuP 184
```

```
        |:::|||   |||   :::||||||::::::|||   :::   :::|||
    644 GTTGGGTGTTGCGGGAGTTCCAGCTATTGTTCAGTTTGTGCTTATGTTGT 693

    184 heLeuProAlaGly..................................... 188
          ||||||   :::
    694 CGCTACCTGAGTCTCCAAGGTGGTTATATAGAAAGGACAGGATTGCGGAA 743

    189 ..............................ThrAspGluThrAlaThrHi 195
                                       :::|||||||:::   :::::
    744 TCCAGAGCGATTCTTGAGAGGATTTACCCGGCTGACGAAGTGGAGGCGGA 793

    195 sLysAspLeuIleProLeuGlnGlyGlyGluAlaProLys......... 208
        :   :::   :::   |||   :::   ||||||   |||
    794 GATGGAAGCTTTGAAACTG...TCAGTGGAGGCAGAGAAAGCAGACGAGG 840

    209 .....LeuGlyProGlyArgProArgTyrSerPheLeuAspLeuPheArg 223
             :::|||                              |||   |||
    841 CAATCATTGGA.........................GATAGCTTCTCT 863

    224 AlaArg.....................AspAsnMetArgGlyArgThrTh 233
        |||:::                          :::|||        ::
    864 GCAAAGCTGAAAGGAGCTTTCGGGAATCCCGTTGTTCGACGTGGACTCGC 913

    233 rValGlyLeuGlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnV 250
        ::::|||:::   :::   :::   ||||||::::::|||      :::|
    914 TGCTGGTATCACGGTGCAAGTGGCTCAGCAGTTTGTGGGGATCAACACTG 963

    250 alLeuCysTyrAlaSerThrIlePheSerSerValGlyPheHisGlyGly 266
        ||:::   |||:::   :::|||   :::   :::|||:::   ::::::
    964 TCATGTACTACAGTCCTTCTATTGTGCAATTTGCTGGTTACGCCTCTAAC 1013

    267 SerSerAlaValLeuAlaSerValGlyLeuGlyAlaValLysValAlaAl 283
        :::::::|||:::      |||:::      ::::::::::::::   ::
   1014 AAGACAGCTATGGCCTTGTCTCTTATTACTTCCGGTTTGAATGCTTTGGG 1063

    283 aThrLeuThrAlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuL 300
        :::::::::::::|||   :::|||||||||   |||||||||   |||:
   1064 TTCTATTGTGAGCATGATGTTTGTCGATCGGTACGGGAGAAGGAAGCTCA 1113

    300 eu..................LeuAlaGlyCysAlaLeuMetAlaLeuSer 310
        ::                  :::   :::|||   ::::::      :::
   1114 TGATCATCTCTATGTTTGGGATCATAGCATGTCTTATCATCTTAGCCACG 1163

    311 ValSerGlyIleGlyLeuValSerPheAlaValProMetAspSerGlyPr 327
        |||   :::   :::   :::      |||      :::|||:::
   1164 GTGTTCTCACAAGCGGCCATC...CACGCCCCCAAGATCGATGCGTTTGA 1210

    327 oSerCysLeuAlaValProAsnAlaThrGlyGlnThrGlyLeuProGlyA 344
         |||         :::||||||||||||   ::::::      |||
   1211 GTCAAGGACGTTTGCACCAAATGCTACTTGTTCAGCATATGCCCCT.... 1256

    344 spSerGlyLeuLeuGlnAspSerSerLeuProProIleProArgThrAsn 360
                 |||      ::::::   ||||||
   1257 ..........CTCGCAGCCGAAAATGCTCCTCCC.............. 1280

    361 GluAspGlnArgGluProIleLeuSerThrAlaLysLysThrLysProHi 377
              :::|||         ::::::      ::::::
   1281 ......TCGAGGTGGAACTGCATGAAGTGTCTCAGGTCCGAATGTGGATT 1324

    377 sProArgSerGly...AspProSerAlaProProArgLeuAlaLeuSerS 393
              ||||||   :::|||   ||||||         ::::::   |
   1325 CTGCGCCAGCGGGGTACAGCCGTACGCACCAGGAGCATGTGTGGTGTTGT 1374

    393 erAlaLeuProGlyProProLeuProAlaArgGlyHisAlaLeuLeuArg 409
        ||                          :::||||||:::::::::::::::
   1375 CAGATGACATGAAAGCAACTTGCAGCTCAAGAGGAAGAACATTCTTCAAA 1424

    410 ......................TrpThrAlaLeuLeuCysLeuMetVa 418
                               :::   |||::::::   |||   ::
   1425 GATGGATGTCCAAGCAAGTTTGGGTTCTTAGCTATAGTGTTTTTGGGGCT 1474

    418 lPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValL 435
        :::::::   :::::::::   |||:::|||   |||   |||:::|||
```

```
     1475 TTACATCGTAGTCTACGCACCAGGTATGGGCACTGTCCCGTGGATCGTCA 1524

      435 euSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCys 451
          ||||||||||||||||||:::::   ||||||   :::   :::::::::
     1525 ACTCTGAGATCTATCCGCTCAGATACAGAGGTCTTGGTGGAGGAATAGCC 1574

      452 AsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAs 468
              ||||||:::::|||||||:::::::|||   ||||||||||::
     1575 GCCGTCTCGAATTGGGTCTCCAACCTGATAGTGAGTGAGAGCTTCCTCTC 1624

      468 pLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeuT 485
          :|||      :::::|||   |||   ||||||||||||||:::::
     1625 ACTCACCCATGCTCTCGGCTCCTCTGGTACCTTCCTTCTCTTTGCTGGCT 1674

      485 hrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLys 501
            ::::::::::||||||   ||||||:::|||:::||||||||||||||
     1675 TCTCCACGATCGGGCTCTTCTTCATTTGGTTGCTCGTTCCTGAGACCAAA 1724

      502 GlyGlnSerLeuAlaGluIleAspGln 510
          |||   ::::::   |||:::::::::
     1725 GGGCTTCAGTTTGAGGAGGTCGAGAAG 1751
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH88793

seq_documentation_block:

```
ID   AAH88793 standard; cDNA; 1905 BP.
XX
AC   AAH88793;
XX
DT   28-SEP-2001  (first entry)
XX
DE   Sugar transporter cDNA sequence #86.
XX
KW   Moss; carbohydrate metabolism related protein; CMRP; sugar; cofactor;
KW   fine chemical production; carbohydrate; polysaccharide; ss.
XX
OS   Physcomitrella patens.
XX
PN   WO200144476-A2.
XX
PD   21-JUN-2001.
XX
PF   14-DEC-2000; 2000WO-EP12697.
XX
PR   16-DEC-1999;   99US-0171101.
XX
PA   (BADI ) BASF PLANT SCI GMBH.
XX
PI   Lerchl J,  Renz A,  Ehrhardt T,  Reindl A,  Cirpus P,  Bischoff F;
PI   Frank M,  Freund A,  Duwenig E,  Schmidt R,  Reski R;
XX
DR   WPI; 2001-398155/42.
DR   P-PSDB; AAM00107.
XX
PT   Novel moss nucleic acid molecules encoding a carbohydrate metabolism
PT   related protein useful for modulating production of fine chemicals such
PT   as carbohydrates, cofactors and enzymes from microorganisms and plants
PT   -
XX
PS   Claim 7; Page 110; 133pp; English.
XX
CC   This invention relates to nucleic acid molecules AAH88708 - AAH88796
CC   isolated from Physcomitrella patens (a moss), which encode carbohydrate
CC   metabolism related proteins (CMRP) represented in AAM00022 - AAM00110.
CC   Included in the invention is a vector containing the CMRP cDNA, and a
CC   host cell transformed with the vector. The host cell (a microorganism,
CC   Corynebacterium or Brevibacterium, moss or algae or a plant cell) is
CC   useful for producing a fine chemical such as carbohydrates, cofactors
CC   and/or enzymes. The nucleic acid molecules are suitable for modifying a
CC   carbohydrate production system in a host, e.g., microorganisms and
CC   plants. They are also useful to identify those DNA sequences and enzymes
CC   in other species which are useful to modify the biosynthesis of starch,
CC   cell wall polysaccharides and sugars. The nucleic acid molecules may be
CC   utilised in the genetic engineering of Corynebacterium glutamicum and the
CC   related Brevibacteium species and Acetobacter xylinum and Chlorella to
CC   make it a better or more efficient producer of one or more fine
CC   chemicals. Mutagenesis of one or more CMRPs may also result in CMRPs
CC   having altered activities which indirectly impact the production of one
CC   or more desired fine chemicals from plants. Primers AAH88705 - AAH80707
CC   are used in the sequencing of the CMRP cDNA sequences of the invention.
XX
SQ   Sequence 1905 BP; 402 A; 408 C; 538 G; 557 T; 0 other;
```

alignment_scores:

```
            Quality:  517.50              Length:    527
              Ratio:  1.719                 Gaps:      8
 Percent Similarity:  57.116   Percent Identity: 28.463
```

alignment_block:

US-09-652-292-2 x AAH88793 ..

Align seg 1/1 to: AAH88793 from: 1 to: 1905

```
       7 ValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLeuThrPheGl 23
         |||||||||   ::::::::::::::::   ||||||::::::   |||||
     106 GTCTTGCCT...GCAGTTGGTATTGCCTGCCTTGGGGCAATCTTGTTTGG 152

      23 yTyrGluLeuAlaValIleSerGlyAlaLeu......LeuProLeuGlnL 38
         ||||:::|||:::|||::::::|||||||||      :::      :::|
     153 TTACCATCTCGGGGTGGTTAATGGTGCATTGGAGTACATTTCTAAGGATC 202

      38 euAspPheGlyLeuSerCysLeuGluGlnGluPheLeuValGlySerLeu 54
         ||   |||:::   ::::::::::::|||   ::::::|||:::|||
     203 TAGGGTTTGCCACGGATGCTGTAAAACAAGGATGGGTGGTAAGCTCAACT 252

      55 LeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLeuIleAspCy 71
         |||   ||||||   ::::::|||::::::||||||   |||   |||
     253 CTAGCTGGTGCCACTGTGGGTTCCTTTACTGGAGGCGCCCTTGCTGACAA 302

      71 sTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValLeuLeuAlaG 88
             |||||||||:::::::::      :::   :::   |||:::::::|
     303 CTTAGGTCGCAAGCGTACATTCCAGATTAACGCCGTGCCTCTTATTGTGG 352

      88 lySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeuValLeuGly 104
         ||:::|||      :::   |||   |||:::      :::|||:::|||
     353 GCACTCTTCTCAGTGCAAAAGCAACCAGTTTCGAGGCTATGGTGATTGGA 402

     105 ArgAlaValValGlyPheAlaIleSerLeuSerSerMetAlaCysCysIl 121
         |||   :::|||||||   :::|||::::::||||||   :::      ::
     403 AGAATTTTGGTTGGTGTTGGGATTGGAGTTTCATCTGGTGTTGTGCCTCT 452

     121 eTyrValSerGluLeuValGlyProArgGlnArgGlyValLeuValSerL 138
         :|||:::||||||:::         :::   ||||||::::::   :::|
     453 ATACATTTCGGAGGTCTCGCCCACAGAGATTCGAGGTACCATGGGGACAT 502

     138 euTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyrAlaLeuAsn 154
         ||   :::      |||   ||||||||||||||||:::      :::
     503 TGAATCAGCTCTTTATTTGCGTGGGTATCCTGTTAGCTCTGATTGCTGGC 552

     155 TyrAlaLeuAlaGlyThrProTrpGlyTrpArgHisMetPheGlyTrpAl 171
             |||:::::::::|||      ||||||   ||||||:::   ||
     553 CTTCCTTTGGGCAGTAACCCTGTCTGGTGGCGCACCATGTTTGCCTTAGC 602

     171 aThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPheLeuPro.... 186
         ||||:::||||||||||||   :::|||::::::   :::   |||
     603 TACAGTTCCTGCCGTTTTGCTGGGTTTAGGCATGGCGTACTGTCCGGAGA 652

     187 ..........................AlaGlyThrAspGluThrAla 193
                                    ::::::      |||||||||
     653 GTCCACGCTGGCTATACAAGAATGGTAAGACCGCAGAGGCGGAAACCGCA 702

     194 ThrHisLys.........................AspLeuIleProLe 201
         :::::::::                         ::::::      ||
     703 GTAAGGAGACTTTGGGGCAAGGCAAAGGTCGAGAGTTCAATGGCAGATTT 752
```

```
 201 uGlnGlyGlyGluAlaProLysLeuGlyProGlyArgProArgTyrSerP 218
     |:::::::::::::::      :::                     |||: .
 753 GAAGGCTAGCAGCGTGGAAACAGTGAAAGGTGACACTCAAGATGCAAGTT 802

 218 heLeuAspLeuPheArgAlaArgAspAsnMetArgGlyArgThrThrVal 234
     ::   :::||||||     |||         |||      :::||||||
 803 GGGGCGAGCTATTTGGCAAAAGA......TACCGTAAAGTTGTCACGGTT 846

 235 GlyLeuGlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnValLe 251
     |||::::::|||.  ||||||||||||:::::|||        |||::
 847 GGAATGGCGCTCTTCCTTTTCCAACAATTTGCCGGAATCAATGCTGTGGT 896

 251 uCysTyrAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerS 268
     :   :::::::::   :::|||   |||:::|||:::         ::::
 897 ATACTTCTCTACTCAGGTTTTCAGGAGTGCTGGCATC.........ACGA 937

 268 erAlaValLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThr 284
     ::   |||   ||||||:::   :::||||||:::::|||||||:::|||
 938 ATGATGTAGCTGCCAGTGCTCTTGTAGGTGCTGCAAATGTGGCAGGTACC 987

 285 LeuThrAlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLe 301
        :::|||   |||:::::|||:::   ||||||:::::||||||::
 988 ACTGTGGCGTCCGGCATGATGGATAAGCAAGGGCGTAAGAGCCTGCTAAT 1037

 301 uAlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValS 318
     :::::::   |||   |||:::||||||:::
1038 GGGCAGCTTCGCTGGCATGTCACTTTCAATG..................... 1068

 318 erPheAlaValProMetAspSerGlyProSerCysLeuAlaValProAsn 334

1068 ..................................................... 1068

 335 AlaThrGlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSe 351

1068 ..................................................... 1068

 351 rSerLeuProProIleProArgThrAsnGluAspGlnArgGluProIleL 368

1068 ..................................................... 1068

 368 euSerThrAlaLysLysThrLysProHisProArgSerGlyAspProSer 384

1068 ..................................................... 1068

 385 AlaProProArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuPr 401
              |||:::||||||   ||||||          ||||||
1069 ...........CTTGTGCTTTCGTTGGCGCTTTCATGGAGCCCCCTTGC 1106

 401 oAlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetV 418
           :::   :::|||            |||:::|||      :::
1107 ACCGTACTCTGGCACGCTG............GCTGTCCTTGGAACAGTTT 1144

 418 alPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuVal 434
       ::::::   :::|||||||:::|||   |||||||||      |||:::
1145 CATACATTTTGTCCTTCTCCCTTGGTGCTGGACCAGTGCCTGGCCTTCTG 1194

 435 LeuSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCy 451
     |||   |||||||:::   ::::::||||||:::::|||   |||:::
1195 TTGCCCGAGATCTTCGGTGCTCGCATCCGTGCTAAGGCCGTTGCTCTTTC 1244

 451 sAsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuA 468
         :::   :::|||   :::|||::::::|||:::|||   ||||||:
1245 TCTCGGTGTCCACTGGATTTGTAACTTCATGATTGGACTATTTTTCTTGA 1294

 468 spLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeu 484
     ::::::::      :::|||:::|||   :::::::|||:::::::::
1295 ACGTCGTTCAGAAGTTCGGTGTTAGCACAGTATATCTCTTCTTCTCTGCA 1344

 485 ThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLy 501
     :::::::::   ::::::::::::::          |||   ||||||||
1345 GTATGCGCGGCAGCAATTGCCTATGTAGGCGGTAATGTGGTAGAAACAAA 1394

 501 sGlyGlnSerLeuAlaGluIleAspGlnGln 511
```

```
          ||||:::|||||||   :::|||:::::::::
     1395 GGGGCGGTCACTGGAGGACATCGAACGCGAG 1425

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ24474

seq_documentation_block:
ID   AAZ24474 standard; cDNA; 1862 BP.
XX
AC   AAZ24474;
XX
DT   17-FEB-2000  (first entry)
XX
DE   Spinach glucose translocator cDNA.
XX
KW   Glucose translocator; spinach; plant; herbicide production;
KW   carbon/nitrogen relationship; starch content; sugar formation;
KW   starch mobilization; plastid; hexose transport; ss.
XX
OS   Spinacia oleracea.
XX
FH   Key             Location/Qualifiers
FT   CDS             61..1716
FT                   /*tag= a
FT                   /product= "glucose translocator"
XX
PN   DE19826444-C1.
XX
PD   18-NOV-1999.
XX
PF   13-JUN-1998;   98DE-1026444.
XX
PR   13-JUN-1998;   98DE-1026444.
XX
PA   (FLUE/) FLUEGGE U.
PA   (SERV/) SERVAITES J.
XX
PI   Fluegge U,  Weber A,  Fischer K,  Servaites J;
XX
DR   WPI; 1999-621474/54.
DR   P-PSDB; AAY50799.
XX
PT   DNA encoding a spinach glucose translocator, plasmids, bacteria, yeast
PT   and transformed plant cells  -
XX
PS   Claim 1; Fig 1; 12pp; German.
XX
CC   This invention describes a novel DNA sequence (I) containing the coding
CC   region of a glucose translocator (I) or a variant encoding a protein
CC   having the biological activity of the glucose translocator which are
CC   useful for identification of insertion mutants, for homologous
CC   recombination or to express a non-translatable RNA, as an agent with an
CC   antisense effect for cosuppression or a ribozyme activity preventing
CC   synthesis of one or more endogenous plastid glucose translocators in the
CC   cell. This is very useful for production of herbicides. The sequences can
CC   be used to alter the carbon/nitrogen relationship in leaves in
CC   heterotrophic tissue, in particular to increase starch content. It can
CC   also be used to degrade sugar formation during starch mobilization. The
CC   sequences are also useful to isolate DNA encoding homologues and
CC   corresponding genomic clones, in particular the corresponding promoter
CC   region or partial promoter region for tissue specific gene expression.
CC   (I) and its variants also serve as a target sequence and with the help of
CC   a prokaryotic or eukaryotic protein sequence, catalyses active or passive
CC   transport of metabolites over membranes, to direct the metabolites into
CC   the plastid membrane, plastid stroma or thylakoids. The sequences can
CC   also be used to code a mature protein with the biological activity of a
CC   glucose translocator, which can be directed to another cell compartment
CC   or cellular membrane system. They can also be used to identify a
CC   substance, which inhibits activity of hexose transport over the inner
CC   plastid membrane. This sequence encodes the spinach glucose translocator
CC   described in the method of the invention.
XX
SQ   Sequence 1862 BP; 437 A; 349 C; 469 G; 607 T; 0 other;
```

```
alignment_scores:
           Quality:  507.00              Length:    525
             Ratio:  1.742                Gaps:     10
 Percent Similarity:  55.429   Percent Identity: 29.714

alignment_block:
 US-09-652-292-2 x AAZ24474    ..

 Align seg 1/1  to: AAZ24474  from: 1  to: 1862

       4 SerProProValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLe 20
         |||      |||||||||      :::::::|||:::  ||||||:::::
     379 TCTGCATCAGTGTTGCCGTAT...GTTGGTGTTGCTTGTTTGGGAGCTAT 425

      20 uThrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeuLeuProLeuG 37
         :  |||||||||:::|||:::|||:::::::|||||||||     |||:
     426 TTTGTTTGGATATCACCTTGGTGTGGTAAATGGAGCTCTTGACTACCTAT 475

      37 lnLeuAspPheGlyLeu......SerCysLeuGluGlnGluPheLeuVal 51
         ::  |||:::::::::      :::   ::   |||   ::::::|||
     476 CTGCGGATTTAGCTATTGCTGGCAATACTGTTTTACAAGGATGGGTTGTG 525

      52 GlySerLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLe 68
         :::   ||||||   ||||||   ::::::|||:::::||||||   ||
     526 AGCATATTACTTGCTGGAGCGACTGTAGGTTCATTTACTGGCGGATCTTT 575

      68 uIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValL 85
         |   |||   :::||||||   :::::::::      :::   :::   |
     576 GGCTGATAAGTTTGGAAGGACAAAGACCTTCCAGTTAGATGCAATTCCTC 625

      85 euLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeu 101
         ||      |||:::          :::   |||  ::::::      :::
     626 TTGCAATAGGAGCTTATCTCTGTGCCACAGCACAAAATGTACAGATAATG 675

     102 ValLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerMetAl 118
         ::::::||||||   :::   |||::::::|||:::::|||||||
     676 ATGATTGGTCGTTTACTTTGTGGCATTGGAATTGGCATATCATCTGCCTT 725

     118 aCysCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValL 135
             :::|||:::||||||||:::          :::   ||||||:::|
     726 AGTTCCGTTATACATTTCTGAGATTTCGCCAACTGAAATTCGAGGTGCAC 775

     135 euValSerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyr 151
         ||   |||:::  :::      |||   :::||||||||||  :::
     776 TTGGTTCTGTCAACCAGCTTTTTATATGCATTGGTATTCTTGCGGCTTTG 825

     152 AlaLeuAsnTyrAlaLeuAlaGlyThrProTrpGlyTrpArgHisMetPh 168
         :::   :::      |||||||||:::|||      ||||||   |||||
     826 GTTGCTGGATTGCCTTTAGCAGGAAACCCTTTGTGGTGGAGGACAATGTT 875

     168 eGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPheL 185
         ||||   ||||||:::|||:::||||||   :::|||::::::   |||
     876 TGGTATTGCAACAGTTCCCTCTGTTCTCTTGGCTCTTGGGATGGGCTTCT 925

     185 euProAlaGlyThr.....................AspGluThrAlaThr 194
           |||   :::                        :::   :::
     926 GTCCAGAAAGCCCTAGGTGGCTTTTTCAGCAAGGAAAAATTGTTGAAGCA 975

     195 HisLysAspLeuIleProLeuGlnGlyGlyGlu.................. 205
         :::|||   :::      |||   |||   |||
     976 GAAAAGGCCGTAGCAGCCCTCTATGGAAAAGAAAGGGTTCCGGAGGTGAT 1025

     206 .............AlaProLysLeuGlyProGlyArgProArgTyrSerP 218
                     |||         |||   ::::::|||:::   ::::
    1026 TAATGATTTGAGAGCATCTGTTCAAGGTTCTTCTGAACCAGAAGCTGGCT 1075

     218 heLeuAspLeuPheArgAlaArgAspAsnMetArgGlyArgThrThrVal 234
         :::::|||||||||   :::|||          |||          :::::|||
    1076 GGTTTGATCTGTTCAGTAGCCGC......TACAGGAAAGTGGTAAGTGTT 1119

     235 GlyLeuGlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnValLe 251
         |||   :::|||   ||||||||||||:::::|||         |||::

    1120 GGTGCTGCACTTTTCTTGTTCCAGCAAATGGCAGGAATCAATGCTGTTGT 1169

     251 uCysTyrAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerS 268
         :  |||:::::::::::|||   |||:::|||:::          :::|
    1170 ATATTACTCCACATCTGTATTCCGAAGTGCTGGAATT.........GCAT 1210

     268 erAlaValLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThr 284
         ||   |||   |||||||:::   :::|||||||   :::|||   :::|||
    1211 CAGATGTTGCCGCTAGTGCCCTTGTTGGAGCATCAAATGTTATTGGCACT 1260

     285 LeuThrAlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLe 301
         :::|||   :::|||:::|||:::   ||||||::::::|||||||::
    1261 GCTGTCGCCTCTTCTTTAATGGACAAGCAAGGAAGAAAGAGCCTACTGAT 1310

     301 uAlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValS 318
         :::::::   :::   ||||||   |||:::      :::::|||   |
    1311 GACAAGTTTCTCAGGAATGGCTGCCTCAATGTTGTTGCTGTCACTA...T 1357

     318 erPheAlaValProMetAspSerGlyProSerCysLeuAlaValProAsn 334
         |||||
    1358 CGTTC..................................................... 1362

     335 AlaThrGlyGlnThrGlyLeuProGlyAspSerGlyLeuGlnAspSe 351

    1362 ..... 1362

     351 rSerLeuProProIleProArgThrAsnGluAspGlnArgGluProIleL 368

    1362 ..... 1362

     368 euSerThrAlaLysLysThrLysProHisProArgSerGlyAspProSer 384
             |||  |||          |||:::      ||||||
    1363 .....ACATGGAAGGCTCTAGCACCTTAC......TCGGGCACA..... 1395

     385 AlaProProArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuPr 401

    1395 ..... 1395

     401 oAlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetV 418
                                        |||::::::      ::::
    1396 ..........................CTTGCCGTTGTGGGAACTGTTT 1417

     418 alPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuVal 434
         ::::::|||   :::||||||:::|||   |||||||||      |||:::
    1418 TATATGTCCTTTCCTTTTCACTTGGCGCTGGTCCTGTTCCTGCTCTTCTA 1467

     435 LeuSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCy 451
         |||   ||||||:::      :::||||||:::::|||   |||:::
    1468 CTTCCAGAGATCTTCGCATCCAGAATCCGAGCAAAAGCAGTGGCTTTGTC 1517

     451 sAsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuA 468
         ::::::::      ||||||:::|||:::   |||:::|||   ||||||:
    1518 ACTGGGAATGCATTGGGCGTCAAACTTCGTAATTGGCCTTTATTTCTTAA 1567

     468 spLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeu 484
         ::::::::      :::|||:::|||   :::::|||   ::::::
    1568 GTGTGGTGACCAAGTTTGGAATCAGCAAAGTGTATTTGGGGTTCGCGTCA 1617

     485 ThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLy 501
         ::::::||||||:::::   :::|||          |||   ||||||||
    1618 GTGTGTGTGCTTGCCGTGTTGTACATAGCTGGTAACGTTGTTGAAACAAA 1667

     501 sGlyGlnSerLeuAlaGluIleAsp 509
         ||||:::|||||||   |||||||::
    1668 AGGACGATCACTTGAGGAAATAGAG 1692

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC36954

seq_documentation_block:
ID   AAC36954 standard; DNA; 2134 BP.
XX
AC   AAC36954;
XX
```

DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15657.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

```
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

alignment_scores:
            Quality:  506.00              Length:    525
              Ratio:  1.745                Gaps:      8
 Percent Similarity:  55.238   Percent Identity: 29.524

alignment_block:
 US-09-652-292-2 x AAC36954   ..

 Align seg 1/1  to: AAC36954  from: 1  to: 2134

       4 SerProProValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLe 20
         |||       |||||||||:::   :::::::|||:::   ||||||:::::
     661 TCTGGAACAGTTTTGCCTTTT...GTTGGTGTTGCTTGTCTTGGTGCTAT 707

      20 uThrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeuLeuProLeuG 37
         :  ||||||||||:::|||::|||::::::|||||||||       |||
     708 ACTCTTTGGTTATCATCTCGGGGTGGTTAATGGTGCTCTTGAATATCTTG 757

      37 lnLeuAspPheGlyLeu......SerCysLeuGluGlnGluPheLeuVal 51
           |||:::|||:::      :::   :::   |||   :::::::|||
     758 CTAAGGATCTTGGGATCGCCGAAAATACTGTTTTGCAAGGATGGATTGTT 807

      52 GlySerLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyPheLe 68
         :::|||||||||   ||||||   ::::::|||::::::||||||   ||
     808 AGTTCTCTGCTTGCTGGTGCTACGGTAGGTTCATTCACTGGAGGTGCATT 857

      68 uIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeuValL 85
         |   |||   :::|||||||   :::::::::       :::   :::   |
     858 AGCTGACAAATTTGGACGAACAAGAACTTTTCAATTGGATGCTATCCCGC 907

      85 euLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTrpLeu 101
         ||      |||:::::::       :::   |||   |||:::      :::
     908 TTGCCATTGGAGCTTTCTTATGTGCAACAGCTCAGAGTGTGCAGACTATG 957

     102 ValLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerMetAl 118
         ::::::||||||   ::::::|||:::::::|||:::::||||||
     958 ATTGTGGGACGTCTGCTCGCTGGAATTGGAATTGGAATCTCATCAGCGAT 1007

     118 aCysCysIleTyrValSerGluLeuValGlyProArgGlnArgGlyValL 135
           :::|||:::|||||||:::          :::   |||||||:::|
    1008 TGTACCACTTTACATATCTGAGATATCACCAACTGAAATCCGTGGAGCAC 1057

     135 euValSerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSerTyr 151
         ||   |||:::   :::       |||   :::|||||||||   :::
    1058 TCGGATCTGTGAACCAGTTGTTCATCTGTATAGGAATACTTGCAGCCTTG 1107

     152 AlaLeuAsnTyrAlaLeuAlaGlyThrProTrpGlyTrpArgHisMetPh 168
           :::       ||||||:::::|||       ||||||   |||||
    1108 ATAGCTGGATTACCCCTTGCAGCAAACCCTCTATGGTGGAGGACGATGTT 1157

     168 eGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeuPheL 185
         ||||   |||:::   |||:::||||||   ::::::::::::   |||
    1158 TGGTGTTGCAGTTATCCCTTCCGTTCTATTGGCCATAGGAATGGCTTTTT 1207

     185 euProAlaGlyThr.....................AspGluThrAlaThr 194
           |||   :::                                 :::   :::
    1208 CTCCAGAAAGCCCAAGGTGGCTCGTTCAGCAAGGAAAAGTCTCTGAAGCT 1257

     195 HisLysAspLeuIleProLeuGlnGlyGlyGlu..................... 205
         :::|||   :::       |||   |||   |||
    1258 GAAAAGGCGATCAAAACTTTGTATGGTAAAGAAAGAGTGGTTGAACTAGT 1307

     206 .............AlaProLysLeuGlyProGlyArgProArgTyrSerP 218
                       |||           |||   ::::::|||:::   ::::
    1308 TCGCGACTTATCAGCCTCTGGCCAAGGTTCTTCTGAGCCGGAGGCAGGAT 1357

     218 heLeuAspLeuPheArgAlaArgAspAsnMetArgGlyArgThrThrVal 234
         :::::||||||||||   :::|||                  ::::::|||
    1358 GGTTTGATCTATTCAGCAGCCGC......TACTGGAAAGTTGTAAGCGTA 1401

     235 GlyLeuGlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnValLe 251
         |||   :::|||   |||||||||||||||:::|||          |||::
    1402 GGTGCGGCTCTCTTCTTGTTTCAACAGTTAGCCGGGATAAACGCAGTTGT 1451

     251 uCysTyrAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerS 268
         :   |||::::::::::|||   |||:::|||:::          :::|
    1452 GTATTACTCCACATCGGTATTCCGTAGTGCGGGAATC.........CAAT 1492

     268 erAlaValLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThr 284
         ||   |||   |||||||:::   :::||||||   :::|||   :::|||
    1493 CAGATGTTGCAGCCAGTGCTCTCGTTGGAGCATCAAATGTCTTTGGCACT 1542

     285 LeuThrAlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLe 301
         :::|||   :::|||:::|||:::   ||||||:::::|||||||||
    1543 GCTGTTGCTTCATCGTTGATGGATAAAATGGGAAGGAAAAGTCTTTTACT 1592

     301 uAlaGlyCysAlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValS 318
         |:::::   :::   |||||||||||||:::
    1593 GACAAGCTTTGGTGGAATGGCTTTGTCAATG..................... 1623

     318 erPheAlaValProMetAspSerGlyProSerCysLeuAlaValProAsn 334

    1623 ..................................................... 1623

     335 AlaThrGlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSe 351
```

```
    1623 ..................................................  1623

     351 rSerLeuProProIleProArgThrAsnGluAspGlnArgGluProIleL  368

    1623 ..................................................  1623

     368 euSerThrAlaLysLysThrLysProHisProArgSerGlyAspProSer  384

    1623 ..................................................  1623

     385 AlaProProArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuPr  401
                  |||   ||||||   ::::::             |||
    1624 ...........CTGTTACTCTCCTTGTCCTTCACATGGAAGGCTCTTGC  1661

     401 oAlaArgGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetV  418
          |||   :::   :::|||           |||::::::     ::::
    1662 TGCCTATTCTGGAACCCTT...........GCCGTTGTTGGAACTGTTC  1699

     418 alPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuVal  434
         :::::|||   :::|||||||:::|||   |||||||||       |||:::
    1700 TATATGTCCTGTCATTCTCACTTGGTGCTGGCCCGGTACCGGCTCTTCTT  1749

     435 LeuSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCy  451
         |||   ||||||:::      :::|||||||:::::|||   |||:::
    1750 CTTCCAGAGATATTTGCATCCCGAATCAGAGCAAAAGCCGTCGCTCTTTC  1799

     451 sAsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuA  468
             :::::::::|||   :::|||:::   |||:::|||   ||||||:
    1800 TCTCGGCATGCACTGGATATCAAACTTTGTGATCGGACTATACTTCTTAA  1849

     468 spLeuIleGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeu  484
         ::::::::      :::|||:::|||   :::::::|||   ::::::
    1850 GCGTTGTGACTAAATTCGGAATCAGCAGTGTCTACTTGGGTTTTGCTGGA  1899

     485 ThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLy  501
         ::::::|||||:::::   :::|||         |||   ||||||||
    1900 GTCTGCGTCCTTGCGGTCCTCTACATTGCAGGAAACGTCGTCGAGACTAA  1949

     501 sGlyGlnSerLeuAlaGluIleAsp  509
         ||||:::||||||   ||||||:::
    1950 AGGTCGATCACTGGAGGAAATAGAG  1974

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA65400

seq_documentation_block:
ID   AAA65400 standard; DNA; 1476 BP.
XX
AC   AAA65400;
XX
DT   09-NOV-2000  (first entry)
XX
DE   Brevibacterium lactofermentum gltBD gene SEQ ID NO:5.
XX
KW   Brevibacterium lactofermentum; gltBD; sugar transporter; breeding;
KW   corynebacterium; phosphoenolpyruvate-sugar transport system; PTS;
KW   non-PTS; ds.
XX
OS   Brevibacterium lactofermentum.
XX
PN   WO200037497-A1.
XX
PD   29-JUN-2000.
XX
PF   16-DEC-1999;   99WO-JP07078.
XX
PR   18-DEC-1998;   98JP-0360620.
XX
PA   (AJIN ) AJINOMOTO CO INC.
XX
PI   Kanno S,  Kimura E,  Matsui K,  Nakamatsu T;
XX
DR   WPI; 2000-442642/38.
DR   P-PSDB; AAB12594.
XX
PT   Sugar transporter gene gltBD of Brevibacterium lactofermentum, useful
PT   for production of coryneform bacteria with altered membrane sugar
PT   transport -
XX
PS   Claim 3; Page 17-21; 26pp; Japanese.
XX
CC   The present invention describes a protein which has transmembrane sugar
CC   transport activity. The protein is specifically a non
CC   phosphoenolpyruvate-sugar transport system (PTS) sugar transporter.
CC   It can be used for in breeding corynebacteria with altered transmembrane
CC   sugar transport. The present represents the Brevibacterium lactofermentum
CC   sugar transporter gltBD gene, which is used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1476 BP; 298 A; 394 C; 372 G; 412 T; 0 other;


alignment_scores:
            Quality:  504.00              Length:    524
              Ratio:  1.647                 Gaps:      9
 Percent Similarity:  58.397   Percent Identity: 26.527

alignment_block:
 US-09-652-292-2 x AAA65400   ..

 Align seg 1/1  to: AAA65400  from: 1  to: 1476

      12 AlaSerValSerLeuLeuGlyGlyLeuThrPheGlyTyrGluLeuAlaVa  28
         |||   |||:::   :::|||||||||   |||||||||:::   :::||
      97 GCACTAGTTGCGACGTTTGGTGGGCTGCTCTTCGGCTACGACACCGGCGT  146

      28 lIleSerGlyAlaLeuLeuProLeuGlnLeuAspPheGlyLeuSerCysL  45
         ||||:::|||||||||   |||:::      :::::::|||||:::::::
     147 CATCAACGGTGCACTCAACCCGATGACCCGCGAGCTCGGACTAACCGCGT  196

      45 euGluGlnGluPheLeuValGlySerLeuLeuLeuGlyAlaLeuLeuAla  61
         ::   :::      :::::::::||||||||||:::|||||      :::
     197 TCACCGAGGGTGTTGTAACTTCTTCCCTGCTGTTTGGTGCAGCAGCTGGT  246

      62 SerLeuValGlyGlyPheLeuIleAspCysTyrGlyArgLysGlnAlaIl  78
         ::::::      |||   :::   |||   :::||||||:::::::||
     247 GCGATGTTTTTCGGTCGCATTTCCGACAACTGGGGTCGCCGGAAAACAAT  296

      78 eLeuGlySerAsnLeuValLeuLeuAlaGlySerLeuThrLeuGlyLeuA  95
         |::::::      :::::::::::::::|||:::::::::      :::|
     297 CATCTCACTTGCAGTAGCTTTCTTTGTCGGCACCATGGTCTGCGTGTTTG  346

      95 laGlySerLeuAlaTrpLeuValLeuGlyArgAlaValValGlyPheAla  111
         ||   |||:::|||   :::|||:::|||||||:::::::::|||:::|||
     347 CTCCATCTTTTGCAGTAATGGTTGTCGGACGTGTGCTTCTTGGACTCGCA  396

     112 IleSerLeuSerSerMetAlaCysCysIleTyrValSerGluLeuValGl  128
         ::::::   :::|||   :::      :::|||:::::::|||||:::
     397 GTTGGTGGCGCTTCCACTGTTGTCCCTGTCTACCTGGCTGAACTTGCTCC  446

     128 yProArgGlnArgGlyValLeuValSerLeuTyrGluAlaGlyIleThrV  145
             :::   ||||||   |||::::::      |||      |||:::|
     447 TTTTGAAATCCGTGGCTCACTGGCTGGCCGTAATGAGTTGATGATTGTTG  496

     145 alGlyIleLeuLeuSerTyrAlaLeuAsnTyrAlaLeuAlaGlyThrPro  161
         |||||   |||   :::::::::::::|||   |||:::   |||:::
     497 TTGGTCAGCTCGCAGCTTTCGTCATCAAT...GCGATTATTGGAAATGTT  543

     162 TrpGly..............TrpArgHisMetPheGlyTrpAlaThrAl  173
         :::|||               |||||||:::|||:::::::   |||:::
     544 TTTGGACACCACGATGGTGTGTGGCGCTACATGCTGGCAATTGCCGCAAT  593

     173 aProAlaValLeuGlnSerLeuSerLeuLeuPheLeuProAlaGly....  188
          |||||||::         ::::::::|||   :::|||   :::
     594 CCCAGCAATTGCCCTCTTCTTTGGAATGCTCCGAGTTCCAGAATCCCCAC  643
```

```
 189 .....ThrAspGluThr...AlaThrHisLys 196
                         |||||||:::   |||:::   :::
 644 GCTGGCTTGTTGAGCGAGGACGCATTGATGAGGCTCGCGCAGTTCTTGAA 693

 197 AspLeuIleProLeuGlnGlyGlyGluAla..... 206
      :::   |||||||:::   ::::::|||
 694 ACCATTCGCCCTCTTGAACGTGCCCATGCAGAAGTTGCTGATGTTGAGCA 743

 207 .....ProLysLeuGlyProGlyArgProArgTyrSerPheLeuA 220
                  :::   :::   ::::::          ::::::   :
 744 CCTAGCAAGAGAAGAGCACGCCGTTTCCGAGAAGTCCATGGGCTTAAGGG 793

 220 spLeuPheArgAlaArgAspAsnMetArgGlyArgThrThrValGlyLeu 236
     ::::::::   ::::::   :::|||          |||||||:::
 794 AAATTTTGTCCAGCAAATGGCTTGTACGC...ATCCTCCTAGTAGGTATC 840

 237 GlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnValLeuCysTy 253
     ||||||   :::   |||||||||||||||      :::::::::   ||
 841 GGATTGGGTGTCGCACAGCAGCTGACCGGCATTAACTCCATCATGTATTA 890

 253 rAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerSerAlaV 270
     |:::::::::::::::   ::::::||||||   :::   ::::::|||:
 891 CGGCCAGGTTGTTCTCATTGAGGCTGGTTTC...TCCGAAAATGCAGCTC 937

 270 alLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThrLeuThr 286
     :::::|||:::|||:::   |||:::::::   |||:::::::::::::
 938 TGATCGCCAACGTGGCACCTGGAGTTATCGCAGTTGTCGGTGCATTCATC 987

 287 AlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLeuAlaGl 303
     |||:::   ::::::||||||   :::|||||||:::   |||::::::||
 988 GCACTGTGGATGATGGATCGTATCAACCGCCGTACCACCCTCATTACCGG 1037

 303 yCysAlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValSerPheA 320
     |   :::|||   ::::::|||          ||||||:::::|||:::|
1038 TTATTCTCTCACCACCATTAGCCACGTATTGATCGGTATCGCATCCATAG 1087

 320 laValProMetAspSerGlyProSerCysLeuAlaValProAsnAlaThr 336
     ||   |||:::
1088 CATTCCCAGTC..... 1098

 337 GlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSerSerLe 353

1098 ..... 1098

 353 uProProIleProArgThrAsnGluAspGlnArgGluProIleLeuSerT 370

1098 ..... 1098

 370 hrAlaLysLysThrLysProHisProArgSerGlyAspProSerAlaPro 386
                              |||||||||      |||
1099 .....GGCGATCCACTTCGCCCA 1116

 387 ProArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuProAlaAr 403

1116 ..... 1116

 403 gGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetValPheV 420
                      ::::::   |||      ::::::|||||||
1117 .....TACGTTATTTTGACCCTAGTTGTGGTCTTCG 1147

 420 alSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValLeuSer 436
     ||:::::::::::      ::::::   :::||||||:::::||||||
1148 TGGGATCCATGCAGACCTTCCTCAACGTAGCCACCTGGGTCATGCTCTCT 1197

 437 GluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCysAsnSe 453
     |||::::::|||:::   :::||||||   |||:::::::::
1198 GAGCTCTTCCCGCTGGCAATGCGAGGTTTCGCAATCGGTATCTCAGTGTT 1247

 453 rPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAspLeuI 470
     |||   |||   ||||||   |||::::::|||   |||      ::::
1248 CTTCCTCTGGATCGCAAACGCGTTCCTCGGATTGTTCTTCCCCACCATCA 1297

 470 leGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeuThrAla 486
```

```
      ::   ::::::|||||||:::   ||||||:::::::::::::      :::
1298 TGGAAGCAGTAGGACTAACCGGAACCTTCTTCATGTTCGCCGGAATCGGT 1347

 487 ValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLysGlyGl 503
     |||::::::::|||   |||||||||       ||||||||||||:::||||::
1348 GTGGTTGCCTTGATCTTCATCTACACCCAGGTTCCTGAAACTCGTGGACG 1397

 503 nSerLeuAlaGluIleAspGln 510
     :::|||   |||||||||:::
1398 TACCTTGGAGGAGATTGATGAG 1419
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH65168

seq_documentation_block:

ID AAH65168 standard; DNA; 1473 BP.
XX
AC AAH65168;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 203.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
DR P-PSDB; AAG89949.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 8; SEQ ID NO: 203; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 1473 BP; 296 A; 388 C; 377 G; 412 T; 0 other;

alignment_scores:

Quality:	502.00	Length:	524
Ratio:	1.651	Gaps:	9
Percent Similarity:	58.015	Percent Identity:	26.527

```
alignment_block:
 US-09-652-292-2 x AAH65168   ..

 Align seg 1/1  to: AAH65168  from: 1  to: 1473

      12 AlaSerValSerLeuLeuGlyGlyLeuThrPheGlyTyrGluLeuAlaVa 28
      97 GCACTAGTTGCGACGTTTGGTGGGCTGCTCTTCGGATATGACACCGGAGT 146

      28 lIleSerGlyAlaLeuLeuProLeuGlnLeuAspPheGlyLeuSerCysL 45
     147 AATCAACGGTGCACTCAACCCAATGACACGTGAGCTCGGACTAACCGCGT 196

      45 euGluGlnGluPheLeuValGlySerLeuLeuLeuGlyAlaLeuLeuAla 61
     197 TCACCGAGGGTGTTGTAACTTCTTCCCTGCTGTTTGGTGCAGCAGCTGGT 246

      62 SerLeuValGlyGlyPheLeuIleAspCysTyrGlyArgLysGlnAlaIl 78
     247 GCGATGTTTTTCGGTCGCATTTCCGACAACTGGGGTCGCCGGAAAACAAT 296

      78 eLeuGlySerAsnLeuValLeuLeuAlaGlySerLeuThrLeuGlyLeuA 95
     297 CATCTCACTTGCAGTAGCTTTCTTTGTCGGCACCATGATCTGCGTGTTTG 346

      95 laGlySerLeuAlaTrpLeuValLeuGlyArgAlaValValGlyPheAla 111
     347 CTCCATCTTTTGCAGTAATGGTTGTCGGACGTGTGCTTCTTGGACTCGCA 396

     112 IleSerLeuSerSerMetAlaCysCysIleTyrValSerGluLeuValGl 128
     397 GTTGGTGGCGCTTCCACTGTTGTCCCTGTCTACCTGGCTGAACTTGCTCC 446

     128 yProArgGlnArgGlyValLeuValSerLeuTyrGluAlaGlyIleThrV 145
     447 TTTTGAAATCCGTGGCTCACTGGCTGGCCGTAATGAGTTGATGATTGTTG 496

     145 alGlyIleLeuLeuSerTyrAlaLeuAsnTyrAlaLeuAlaGlyThrPro 161
     497 TTGGTCAGCTCGCAGCTTTTGTCATCAAT...GCGATTATTGGAAATGTT 543

     162 TrpGly..............TrpArgHisMetPheGlyTrpAlaThrAl 173
     544 TTTGGACACCACGATGGTGTGTGGCGCTACATGCTGGCAATTGCCGCAAT 593

     173 aProAlaValLeuGlnSerLeuSerLeuLeuPheLeuProAlaGly.... 188
     594 CCCAGCAATTGCCCTCTTCTTTGGAATGCTCCGAGTTCCAGAATCCCCAC 643

     189 .....ThrAspGluThr...AlaThrHisLys 196
     644 GCTGGCTTGTTGAGCGAGGACGCATTGATGAGGCTCGCGCAGTTCTTGAA 693

     197 AspLeuIleProLeuGlnGlyGlyGluAla..... 206
     694 ACCATTCGCCCTCTAGAACGTGCCCATGCAGAAGTTGCTGATGTTGAACA 743

     207 .........ProLysLeuGlyProGlyArgProArgTyrSerPheLeuA 220
     744 CCTAGCAAGAGAAGAGCACGCCGTTTCCGAGAAGTCCATGGGCTTAAGGG 793

     220 spLeuPheArgAlaArgAspAsnMetArgGlyArgThrThrValGlyLeu 236
     794 AAATTTTGTCCAGCAAGTGGCTTGTGCGC...ATCCTCCTGGTAGGTATC 840

     237 GlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnValLeuCysTy 253
     841 GGATTGGGTGTCGCACAGCAGCTGACCGGCATCAACTCCATCATGTACTA 890

     253 rAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerSerAlaV 270
     891 CGGCCAGGTTGTTCTCATTGAGGCTGGTTTC...TCCGAGAATGCAGCTC 937

     270 alLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThrLeuThr 286
     938 TGATCGCCAACGTGGCGCCAGGAGTGATCGCAGTTGTCGGTGCATTCATC 987

     287 AlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLeuAlaGl 303
     988 GCACTGTGGATGATGGATCGTATCAACCGCCGTACCACCCTCATTACCGG 1037

     303 yCysAlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValSerPheA 320
    1038 TTATTCTCTCACCACCATTAGCCACGTATTGATCGGTATCGCATCCGTAG 1087

     320 laValProMetAspSerGlyProSerCysLeuAlaValProAsnAlaThr 336
    1088 CATTCCCAGTC..... 1098

     337 GlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSerSerLe 353
    1098 ..... 1098

     353 uProProIleProArgThrAsnGluAspGlnArgGluProIleLeuSerT 370
    1098 ..... 1098

     370 hrAlaLysLysThrLysProHisProArgSerGlyAspProSerAlaPro 386
    1099 .....GGCGATCCTCTTCGCCCC 1116

     387 ProArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuProAlaAr 403
    1116 ..... 1116

     403 gGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetValPheV 420
    1117 .....TACGTTATCTTGACTCTGGTTGTGGTCTTCG 1147

     420 alSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValLeuSer 436
    1148 TGGGATCCATGCAGACCTTCCTCAACGTAGCTACCTGGGTTATGCTCTCT 1197

     437 GluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCysAsnSe 453
    1198 GAGCTCTTCCCGCTGGCAATGCGCGGTTTCGCAATCGGTATCTCAGTGTT 1247

     453 rPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAspLeuI 470
    1248 CTTCCTCTGGATCGCAAACGCGTTCCTCGGATTGTTCTTCCCAACCATCA 1297

     470 leGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeuThrAla 486
    1298 TGGAAGCAGTAGGACTAACCGGAACCTTCTTCATGTTCGCCGGAATCGGT 1347

     487 ValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLysGlyGl 503
    1348 GTGGTTGCCTTGATCTTCATCTACACCCAGGTTCCTGAAACTCGTGGACG 1397

     503 nSerLeuAlaGluIleAspGln 510
    1398 TACCTTGGAGGAGATTGATGAG 1419

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH64966

seq_documentation_block:
ID   AAH64966 standard; DNA; 349980 BP.
XX
AC   AAH64966;
XX
DT   26-SEP-2001  (first entry)
XX
DE   C glutamicum coding sequence fragment SEQ ID NO: 1.
XX
KW   Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
```

KW organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT -
XX
PS Claim 7; SEQ ID NO: 1; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 other;

```
alignment_scores:
            Quality:  502.00              Length:    524
              Ratio:  1.651                Gaps:      9
Percent Similarity:  58.015   Percent Identity: 26.527

alignment_block:
US-09-652-292-2 x AAH64966/rev   ..

Align seg 1/1  to reverse of: AAH64966  from: 1  to: 349980

     12 AlaSerValSerLeuLeuGlyGlyLeuThrPheGlyTyrGluLeuAlaVa 28
        |||   |||:::   :::|||||||||  |||||||||:::   :::||
 192079 GCACTAGTTGCGACGTTTGGTGGGCTGCTCTTCGGATATGACACCGGAGT 192030

     28 lIleSerGlyAlaLeuLeuProLeuGlnLeuAspPheGlyLeuSerCysL 45
        ||||:::|||||||||  |||:::      ::::::||||||:::::::
 192029 AATCAACGGTGCACTCAACCCAATGACACGTGAGCTCGGACTAACCGCGT 191980

     45 euGluGlnGluPheLeuValGlySerLeuLeuLeuGlyAlaLeuLeuAla 61
        ::   :::      ::::::::::|||||||||:::||||||     :::
 191979 TCACCGAGGGTGTTGTAACTTCTTCCCTGCTGTTTGGTGCAGCAGCTGGT 191930

     62 SerLeuValGlyGlyPheLeuIleAspCysTyrGlyArgLysGlnAlaIl 78
        ::::::      |||   :::   |||   :::||||||:::::::::||
 191929 GCGATGTTTTTCGGTCGCATTTCCGACAACTGGGGTCGCCGGAAAACAAT 191880

     78 eLeuGlySerAsnLeuValLeuLeuAlaGlySerLeuThrLeuGlyLeuA 95
        |::::::     ::::::::::::::|||::::::         :::|

 191879 CATCTCACTTGCAGTAGCTTTCTTTGTCGGCACCATGATCTGCGTGTTTG 191830

     95 laGlySerLeuAlaTrpLeuValLeuGlyArgAlaValValGlyPheAla 111
        ||   |||:::|||   :::|||:::||||||:::::::::|||:::|||
 191829 CTCCATCTTTTGCAGTAATGGTTGTCGGACGTGTGCTTCTTGGACTCGCA 191780

    112 IleSerLeuSerSerMetAlaCysCysIleTyrValSerGluLeuValGl 128
        ::::::   :::|||   :::      :::|||::::::||||||:::
 191779 GTTGGTGGCGCTTCCACTGTTGTCCCTGTCTACCTGGCTGAACTTGCTCC 191730

    128 yProArgGlnArgGlyValLeuValSerLeuTyrGluAlaGlyIleThrV 145
        :::  ||||||  |||:::::      |||      |||:::|
 191729 TTTTGAAATCCGTGGCTCACTGGCTGGCCGTAATGAGTTGATGATTGTTG 191680

    145 alGlyIleLeuLeuSerTyrAlaLeuAsnTyrAlaLeuAlaGlyThrPro 161
        |||||   |||   :::::::::::|||   |||:::   |||:::
 191679 TTGGTCAGCTCGCAGCTTTTGTCATCAAT...GCGATTATTGGAAATGTT 191633

    162 TrpGly.............TrpArgHisMetPheGlyTrpAlaThrAl 173
        :::|||                |||||:::|||:::::   |||:::
 191632 TTTGGACACCACGATGGTGTGGCGCTACATGCTGGCAATTGCCGCAAT 191583

    173 aProAlaValLeuGlnSerLeuSerLeuLeuPheLeuProAlaGly.... 188
        ||||||:::          ::::::::|||   :::|||   :::
 191582 CCCAGCAATTGCCCTCTTCTTTGGAATGCTCCGAGTTCCAGAATCCCCAC 191533

    189 .....ThrAspGluThr...AlaThrHisLys 196
                           ||||||||:::   |||:::   :::
 191532 GCTGGCTTGTTGAGCGAGGACGCATTGATGAGGCTCGCGCAGTTCTTGAA 191483

    197 AspLeuIleProLeuGlnGlyGlyGluAla..... 206
        :::   ||||||:::   :::::::|||
 191482 ACCATTCGCCCTCTAGAACGTGCCCATGCAGAAGTTGCTGATGTTGAACA 191433

    207 .....ProLysLeuGlyProGlyArgProArgTyrSerPheLeuA 220
                :::   :::   ::::::          ::::::   :
 191432 CCTAGCAAGAGAAGAGCACGCCGTTTCCGAGAAGTCCATGGGCTTAAGGG 191383

    220 spLeuPheArgAlaArgAspAsnMetArgGlyArgThrThrValGlyLeu 236
        ::::::::  ::::::      :::|||            ||||||:::
 191382 AAATTTTGTCCAGCAAGTGGCTTGTGCGC...ATCCTCCTGGTAGGTATC 191336

    237 GlyLeuValLeuPheGlnGlnLeuThrGlyGlnProAsnValLeuCysTy 253
        ||||||   :::   ||||||||||||||      :::::::::   ||
 191335 GGATTGGGTGTCGCACAGCAGCTGACCGGCATCAACTCCATCATGTACTA 191286

    253 rAlaSerThrIlePheSerSerValGlyPheHisGlyGlySerSerAlaV 270
        |:::::::::::::::   ::::::||||||   :::   ::::::|||:
 191285 CGGCCAGGTTGTTCTCATTGAGGCTGGTTTC...TCCGAGAATGCAGCTC 191239

    270 alLeuAlaSerValGlyLeuGlyAlaValLysValAlaAlaThrLeuThr 286
        :::::|||:::|||:::   |||::::::   |||:::::::::::::
 191238 TGATCGCCAACGTGGCGCCAGGAGTGATCGCAGTTGTCGGTGCATTCATC 191189

    287 AlaMetGlyLeuValAspArgAlaGlyArgArgAlaLeuLeuLeuAlaGl 303
        |||:::   :::::::|||||   :::|||||||:::   |||:::::||
 191188 GCACTGTGGATGATGGATCGTATCAACCGCCGTACCACCCTCATTACCGG 191139

    303 yCysAlaLeuMetAlaLeuSerValSerGlyIleGlyLeuValSerPheA 320
        |   :::|||   :::::::|||         ||||||:::::::|||   |
 191138 TTATTCTCTCACCACCATTAGCCACGTATTGATCGGTATCGCATCCGTAG 191089

    320 laValProMetAspSerGlyProSerCysLeuAlaValProAsnAlaThr 336
        ||   |||:::
 191088 CATTCCCAGTC..... 191078

    337 GlyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSerSerLe 353

 191078 ..... 191078

    353 uProProIleProArgThrAsnGluAspGlnArgGluProIleLeuSerT 370

 191078 ..... 191078
```

```
    370 hrAlaLysLysThrLysProHisProArgSerGlyAspProSerAlaPro 386
                                    |||||||||      |||
 191077 ..............................GGCGATCCTCTTCGCCCC 191060

    387 ProArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuProAlaAr 403

 191060 ................................................. 191060

    403 gGlyHisAlaLeuLeuArgTrpThrAlaLeuLeuCysLeuMetValPheV 420
                        ::::::   |||      ::::::|||||||
 191059 ..................TACGTTATCTTGACTCTGGTTGTGGTCTTCG 191029

    420 alSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValLeuSer 436
        ||::::::::::::      ::::::   :::||||||::::::||||||
 191028 TGGGATCCATGCAGACCTTCCTCAACGTAGCTACCTGGGTTATGCTCTCT 190979

    437 GluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCysAsnSe 453
        |||::::::|||:::   :::||||||   |||:::::::::
 190978 GAGCTCTTCCCGCTGGCAATGCGCGGTTTCGCAATCGGTATCTCAGTGTT 190929

    453 rPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAspLeuI 470
         |||   |||   ||||||   |||::::::|||   |||      ::::
 190928 CTTCCTCTGGATCGCAAACGCGTTCCTCGGATTGTTCTTCCCAACCATCA 190879

    470 leGlyThrIleGlyLeuSerTrpThrPheLeuLeuTyrGlyLeuThrAla 486
        ::   ::::::||||||:::   ||||||:::::::::::::      :::
 190878 TGGAAGCAGTAGGACTAACCGGAACCTTCTTCATGTTCGCCGGAATCGGT 190829

    487 ValLeuGlyLeuGlyPheIleTyrLeuPheValProGluThrLysGlyGl 503
        |||::::::|||   |||||||||      ||||||||||||:::|||::
 190828 GTGGTTGCCTTGATCTTCATCTACACCCAGGTTCCTGAAACTCGTGGACG 190779

    503 nSerLeuAlaGluIleAspGln 510
        ::::|||   |||||||||:::
 190778 TACCTTGGAGGAGATTGATGAG 190757

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:AAZ32202

seq_documentation_block:
ID   AAZ32202 standard; cDNA; 1960 BP.
XX
AC   AAZ32202;
XX
DT   14-JAN-2000  (first entry)
XX
DE   Soybean hexose carrier protein encoding cDNA.
XX
KW   Hexose carrier protein; corn; rice; sorghum; soybean; wheat;
KW   carbohydrate transport; plant carbon partitioning; manipulation;
KW   carbohydrate distribution; ss.
XX
OS   Glycine max.
XX
PN   WO9953082-A2.
XX
PD   21-OCT-1999.
XX
PF   07-APR-1999;   99WO-US07561.
XX
PR   09-APR-1998;   98US-0081131.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Allen SM,  Lightner JE,  Rafalski JA,  Thorpe CJ;
XX
DR   WPI; 1999-620438/53.
DR   P-PSDB; AAY40632.
XX
PT   New hexose carrier proteins used to manipulate carbohydrate transport
PT   -
XX
PS   Claim 2; Page 46-47; 60pp; English.
XX
CC   The present sequence encodes a hexose carrier protein from the
CC   invention, which describes hexose carrier proteins isolated from sorghum,
CC   rice, wheat, soybean and corn. Also describes are: (1) a chimeric gene
CC   comprising a hexose carrier protein polynucleotide operably linked to
CC   regulatory sequences; (2) a transformed host cell comprising the chimeric
CC   gene; and (3) altering the level of expression of a hexose carrier
CC   protein in a host cell, comprising transforming a host cell with the
CC   chimeric gene and growing the cell under expression conditions. Hexose
CC   carrier proteins may be used to manipulate carbohydrate transport and to
CC   alter whole plant carbon partitioning or to manipulate carbohydrate
CC   distribution between cellular compartments.
XX
SQ   Sequence 1960 BP; 481 A; 413 C; 456 G; 601 T; 9 other;


alignment_scores:
            Quality:  500.50              Length:    583
              Ratio:  1.657                 Gaps:     14
 Percent Similarity:  51.801   Percent Identity: 27.616

alignment_block:
 US-09-652-292-2 x AAZ32202   ..

 Align seg 1/1  to: AAZ32202  from: 1  to: 1960

       2 GlyHisSerProProValLeuProLeuCysAlaSerValSerLeuLeuGl 18
         |||          |||   :::   :::   :::   |||:::   :::||
     198 GGAAGCCTCACTCCTTTTGTGACGGTAACATGTATAGTTGCAGCCATGGG 247

      18 yGlyLeuThrPheGlyTyrGluLeuAlaValIleSerGly......AlaL 33
         |||||||   |||||||||:::::::::::::   :::|||      ::::
     248 TGGGTTAATCTTCGGTTACGATATCGGAATTTCAGGTGGGGTGACATCCA 297

      33 euLeuProLeuGlnLeuAspPheGlyLeuSerCysLeuGluGlnGlu... 48
         ::   |||:::   |||   |||      |||   ::::::::::::
     298 TGGATCCGTTTCTGCTCAAGTTTTTCCCGTCGGTGTTCCGGAAGAAGAAT 347

      48 ..... 48

     348 TCCGACAAAACGGTGAACCAGTACTGTCAATACGACAGTCAGACACTGAC 397

      49 .PheLeuValGlySerLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValG 65
          :::::::::::|||||   |||:::|||||||||:::|||||||||:
     398 GATGTTCACGTCGTCGCTGTATCTCGCCGCGTTGCTGTCGTCGTTGGTTG 447

      65 lyGlyPheLeuIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySer 81
         :::::   :::         :::|||||||||   :::::::|||   :::
     448 CCTCCACCGTCACACGTAGGTTCGGCCGRAAAMTCTCCATGCTTTTCGGA 497

      82 AsnLeuValLeuLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLe 98
         :::|||:::::::|||:::|||:::|||      |||:::|||      ::
     498 GGCTTGCTTTTCCTCGNCGGTGCCCTTATCAACGGNTTTGCCCANCACGT 547

      98 uAlaTrpLeuValLeuGlyArgAlaValValGlyPheAlaIleSerLeuS 115
         :     |||::::::|||||   ::::::||||||:::|||   ::::
     548 GTGGNTGCTCATCGTGGGTCGGATCTTGCTCGGGTTCGGTATCCGGTTTG 597

     115 erSerMetAlaCysCysIleTyrValSerGluLeuValGlyProArgGln 131
         ::::::::::      :::   :::||||||::::::      :::
     598 CCAATCAGTCTGTGCCACTCTANCTATCTGAAATGGCTCCATACAAATAT 647

     132 ArgGlyValLeuValSerLeuTyrGluAlaGlyIleThrValGlyIleLe 148
         ||||||:::|||         ::::::   :::||||||||||||||||
     648 AGAGGAGCATTGAACATTGGCTTTCAGTTGTCCATCACTGTTGGTATCCT 697

     148 uLeuSerTyrAlaLeuAsnTyrAlaLeuAlaGlyThrPro......TrpG 163
         |:::::::   :::|||||||||   :::|||                ||||
     698 TGTGGCCAATGTGTTGAACTATTTCTTTGCTAAAATCAAAGGTGGTTGGG 747

     163 lyTrpArgHisMetPheGlyTrpAlaThrAlaProAlaValLeuGlnSer 179
         ||||||||      :::|||   |||   :::||||||::::::   :::
     748 GATGGAGGTTGAGTTTGGGAGGTGCTATGGTCCCTGCCCTTATAATCACA 797
```

```
 180 LeuSerLeuLeuPheLeuProAlaGlyThrAspGluThrAlaThrHisLy 196
 798 GTAGGATCACTAGTCCTTCCA.................GACACTCCCAA 829

 196 sAspLeuIleProLeuGlnGlyGlyGluAlaProLysLeuGlyProGlyA 213
 830 TTCCATGATTGAAAGGGGTGATCGCGAGAAGGCCAAGGCTCAGCTTCAGA 879

 213 rgProArg.................TyrSerPheLeuAspLeuPheArg 223
 880 GAATTCGCGGCATCGACAATGTTGATGAAGAGTTCAATGACCTTGTGGCA 929

 224 AlaArgAspAsn..................................... 227
 930 GCAAGTGAATCCTCTAGCCAAGTGGAGCACCCTTGGAGGAACTTGTTGCA 979

 228 .......MetArgGlyArgThrThrValGlyLeuGlyLeuValLeuPheG 242
 980 AAGAAAGTACAGACCCCACCTCACCATGGCAGTGTTGATTCCATTCTTCC 1029

 242 lnGlnLeuThrGlyGlnProAsnValLeuCysTyrAlaSerThrIlePhe 258
1030 AGCAACTCACTGGAATCAATGTCATCATGTTTTACGCGCCGGTCCTGTTC 1079

 259 SerSerValGlyPheHisGlyGlySerSerAlaValLeuAlaSerValGl 275
1080 AGCTCCATCGGGTTTAAG...GATGATGCTGCTCTAATGTCAGCTGTGAT 1126

 275 yLeuGlyAlaValLysValAlaAlaThrLeuThrAlaMetGlyLeuValA 292
1127 CACCGGCGTTGTTAATGTTGTCGCAACTTGTGTCTCAATTTATGGTGTTG 1176

 292 spArgAlaGlyArgArgAlaLeuLeuLeuAlaGlyCysAlaLeuMetAla 308
1177 ACAAGTGGGGTAGGAGAGCCCTTTTCCTTGAAGGTGGAGTCCAAATGCTC 1226

 309 LeuSerValSerGlyIleGlyLeuValSerPheAlaValProMetAspSe 325
1227 ATT.......................................... 1229

 325 rGlyProSerCysLeuAlaValProAsnAlaThr.............G 337
1230 ..........TGCCAGGCTGTAGTTGCAGCTGCAATTGGAGCAAAGTTTG 1269

 337 lyGlnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSerSerLeu 353
1270 GAACTGATGGGAACCCAGGTGAT..................... 1292

 354 ProProIleProArgThrAsnGluAspGlnArgGluProIleLeuSerTh 370
1292 ..................................................... 1292

 370 rAlaLysLysThrLysProHisProArgSerGlyAspProSerAlaProP 387
1292 ..................................................... 1292

 387 roArgLeuAlaLeuSerSerAlaLeuProGlyProProLeuProAlaArg 403
1292 ..................................................... 1292

 404 GlyHisAlaLeuLeuArgTrpThrAlaLeu.........LeuCysLeuMe 417
1293 .........TTGCCAAAGTGGTATGCAATTGTTGTGGTTCTCTTCATTTG 1333

 417 tValPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuV 434
1334 CATTTATGTATCAGCATTTGCCTGGTCATGGGGTCCCCTAGGTTGGTTGG 1383

 434 alLeuSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPhe 450
1384 TGCCTAGTGAGATCTTTCCCTTGGAGATTCGTTCAGCTGCTCAGAGTATC 1433

 451 CysAsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLe 467
1434 AATGTGTCGGTGAACATGCTTTTCACTTTCTTGATTGCACAAGTCTTCTT 1483

 467 uAspLeuIle......GlyThrIleGlyLeuSerTrpThrPheLeuLeuT 482
1484 GACAATGCTTTGCCACATGAAGTTCGGCTTG.........TTCCTCTTCT 1524

 482 yrGlyLeuThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValPro 498
1525 TTGCCTTCTTCGTGTTGATCATGACATTCTTCGTCTACTTCTTCTTGCCC 1574

 499 GluThrLysGlyGlnSerLeuAlaGluIleAspGlnGlnPheGlnLysAr 515
1575 GAAACAAAGGGCATTCCAATTGAAGAAATGGGGCAGGTTTGGCAGGCACA 1624

 515 gArgPheThrLeuSerPhe...GlyHisArgGlnAsnSerThrGlyIle 530
1625 CCCCTTCTGGTCCAGATTCGTGGAGCATGATGATTATGGCAATGGTGTT 1673
```

```
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC50883

seq_documentation_block:
ID   AAC50883 standard; DNA; 1835 BP.
XX
AC   AAC50883;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 66487.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
```

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
Quality: 488.50 Length: 592
Ratio: 1.561 Gaps: 17
Percent Similarity: 52.872 Percent Identity: 26.182

alignment_block:
US-09-652-292-2 x AAC50883 ..

Align seg 1/1 to: AAC50883 from: 1 to: 1835

```
  4 SerProProValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLe 20
186 ACTCCCTTTGTTCTCTTCACTTGCGTT...GTTGCTGCCATGGGCGGTCT 232

 20 uThrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeu...LeuPro. 35
233 CATCTTCGGATACGATATCGGAATCTCCGGTGGTGTGACGTCTATGCCGT 282

 36 .....LeuGlnLeuAspPheGlyLeuSerCysLeuGluGlnGlu..... 48
283 CTTTCCTCAAGCGATTCTTCCCGTCGGTGTATCGGAAACAACAAGAGGAC 332

 49 .....PheLe 50
333 GCGTCAACGAACCAGTACTGTCAGTACGATAGCCCGACGCTAACGATGTT 382

 50 uValGlySerLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyP 67
383 CACATCGTCTCTATATCTAGCGGCGCTAATTTCGTCGCTGGTGGCTTCCA 432

 67 heLeuIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeu 83
433 CCGTGACAAGAAAGTTCGGACGGCGGCTCTCGATGCTCTTCGGCGGCATA 482

 84 ValLeuLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTr 100
483 CTCTTCTGCGCCGGAGCTCTCATCAATGGTTTCGCCAAACATGTTTGGAT 532

100 pLeuValLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerM 117
533 GCTCATCGTCGGTCGTATCTTGCTTGGTTTCGGTATCGGTTTCGCTAATC 582

117 etAlaCysCysIleTyrValSerGluLeuValGlyProArgGlnArgGly 133
583 AGGCTGTGCCACTGTACCTCTCTGAGATGGCTCCATACAAATACAGAGGA 632

134 ValLeuValSerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSe 150
633 GCTTTAAACATTGGTTTCCAGCTCTCAATTACAATCGGAATCCTCGTCGC 682

150 rTyrAlaLeuAsnTyrAlaLeuAlaGlyThrPro......TrpGlyTrpA 165
683 CGAAGTGCTAAACTACTTCTTCGCCAAGATCAAAGGCGGTTGGGGATGGC 732

165 rgHisMetPheGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSer 181
733 GGCTCAGTCTCGGAGGCGCGGTGGTTCCTGCCTTGATCATAACCATCGGC 782

182 LeuLeuPheLeuPro..... 186
783 TCCCTCGTCCTCCCTGACACTCCCAATTCAATGATCGAGCGTGGCCAACA 832

187 .....AlaGlyThrAspGluThrA 193
833 CGAAGAAGCCAAAACCAAGCTCAGACGAATCCGTGGTGTCGATGACGTCA 882

193 laThrHisLysAspLeuIleProLeuGlnGlyGlyGluAlaProLysLeu 209
883 GCCAAGAGTTTGACGATTTGGTCGCCGCTAGTAAAGAGTCGCAGTCGATA 932

210 GlyProGlyArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAs 226
933 .............GAGCACCCGTGGAGAAACCTCCTCCGCCGCAAG.. 965

226 pAsnMetArgGlyArgThrThrValGlyLeuGlyLeuValLeuPheGlnG 243
966 ....TACCGACCACATCTCACAATGGCCGTTATGATTCCGTTCTTTCAAC 1011

243 lnLeuThrGlyGlnProAsnValLeuCysTyrAlaSerThrIlePheSer 259
1012 AGCTAACCGGAATCAATGTGATTATGTTTTACGCTCCGGTTTTGTTCAAC 1061

260 SerValGlyPheHisGlyGlySerSerAlaValLeuAlaSerValGlyLe 276
1062 ACCATTGGTTTCACGACCGATGCTTCT...CTCATGTCCGCTGTGGTCAC 1108

276 uGlyAlaValLysValAlaAlaThrLeuThrAlaMetGlyLeuValAspA 293
1109 TGGCTCGGTTAACGTGGCCGCTACGCTTGTTTCTATCTACGGTGTTGACA 1158

293 rgAlaGlyArgArgAlaLeuLeuLeuAlaGlyCysAlaLeuMetAlaLeu 309
1159 GATGGGGACGTCGGTTTCTCTTTCTTGAAGGTGGTACACAAATGCTTATA 1208

310 SerValSerGlyIleGlyLeuValSerPheAlaValProMetAspSerGl 326
1208 ..... 1208

326 yProSerCysLeuAlaValProAsnAlaThr.....GlyG 338
1209 .......TGCCAGGCTGTGGTTGCAGCTTGCATAGGGGCCAAGTTTGGGG 1251

338 lnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSerSerLeuPro 354
1252 TAGACGGGACCCCTGGTGAG..... 1271

355 ProIleProArgThrAsnGluAspGlnArgGluProIleLeuSerThrAl 371
1271 ..... 1271

371 aLysLysThrLysProHisProArgSerGlyAspProSerAlaProProA 388
1271 ..... 1271

388 rgLeuAlaLeuSerSerAlaLeuProGlyProProLeuProAlaArgGly 404
1271 ..... 1271

405 HisAlaLeuLeuArgTrpThrAlaLeuLeuCys.....LeuMetVa 418
1272 ......CTACCAAAGTGGTATGCTATAGTGGTTGTAACGTTCATTTGCAT 1315

418 lPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValL 435
1316 CTATGTGGCGGGTTTTGCGTGGTCGTGGGGCCCACTAGGGTGGTTAGTAC 1365

435 euSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCys 451
1366 CGAGTGAAATCTTCCCGTTGGAGATAAGGTCGGCGGCGCAGAGTATCACC 1415

452 AsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAs 468
1416 GTGTCCGTGAACATGATCTTCACGTTCATTATCGCGCAAATCTTCTTGAC 1465

468 pLeuIleGlyThrIle......GlyLeuSerTrpThrPheLeuLeuTyrG 483
1466 GATGCTTTGTCATTTGAAGTTTGGGTTA.....TTCCTTGTTTTCG 1506

483 lyLeuThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGlu 499
1507 CCTTTTTCGTGGTGGTGATGTCGATCTTTGTATACATTTTCTTGCCGGAG 1556
```

```
     500 ThrLysGlyGlnSerLeuAlaGluIleAspGlnGlnPheGlnLysArg.. 515
     1557 ACGAAAGGGATTCCGATAGAGGAGATGGGTCAAGTGTGGAGGTCACACTG 1606

      516 .....ArgPhe.....ThrLeuSerP 521
     1607 GTATTGGTCAAGGTTTGTGGAGGATGGTGAGTATGGGAATGCGCTTGAGA 1656

      521 heGlyHisArgGlnAsnSerThrGly 529
     1657 TGGGCAAGAACAGTAACCAAGCTGGA 1682

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF55866

seq_documentation_block:
ID   AAF55866 standard; cDNA; 2087 BP.
XX
AC   AAF55866;
XX
DT   17-APR-2001  (first entry)
XX
DE   Rat GLUTX1 coding sequence.
XX
KW   Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
KW   hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;
KW   hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS   Rattus sp.
XX
PN   WO200104145-A2.
XX
PD   18-JAN-2001.
XX
PF   14-JUL-2000; 2000WO-IB01042.
XX
PR   14-JUL-1999;   99US-0143907.
PR   27-AUG-1999;   99US-0151140.
PR   23-FEB-2000; 2000US-0184285.
PR   13-JUL-2000; 2000US-0616132.
XX
PA   (UYLA-) UNIV LAUSANNE.
XX
PI   Thorens B,  Ibberson M,  Uldry M;
XX
DR   WPI; 2001-112615/12.
DR   P-PSDB; AAB66933.
XX
PT   Nucleic acids encoding GLUTX glucose transporter proteins, useful in
PT   the prevention, diagnosis and treatment of hexose transport disorders,
PT   e.g. ischemia and diabetes -
XX
PS   Claim 3; Page 71-73; 124pp; English.
XX
CC   The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC   AAB66932-AAB66941). The GLUTX proteins are related to the facultative
CC   glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC   function. The GLUTX proteins may be used in the diagnosis, prevention and
CC   treatment of hexose transport disorders such as ischaemia, diabetes,
CC   hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC   neurodegenerative disease. The present sequence is the coding sequence
CC   for rat GLUTX1.
XX
SQ   Sequence 2087 BP; 355 A; 673 C; 580 G; 479 T; 0 other;


alignment_scores:
            Quality:  486.00            Length:    557
              Ratio:  1.659               Gaps:     17
 Percent Similarity:  52.603   Percent Identity: 27.289

alignment_block:
 US-09-652-292-2 x AAF55866   ..
```

```
Align seg 1/1  to: AAF55866  from: 1  to: 2087

       5 ProProValLeuProLeuCys.AlaSerValSerLeuLeuGlyGlyLeuT 21
      98 CCGCCGGGTCTTCCT...TGCCACCTTCGCCGCCGCCCTGGGCCCGCTCA 144

      21 hrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeuLeuProLeuGln 37
     145 GCTTCGGCTTCGCGCTTGGCTACAGCTCCCCCGCCATCCCCAGCCTGCGG 194

      38 LeuAsp.....PheGlyLeuSerCysLeuGluGlnGluPheLe 50
     195 CGCACCGCACCCCCTGCCCTACGCCTCGGAGACACTGCGGCCTCCTGGTT 244

      50 uValGlySerLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyP 67
     245 CGGGGCCGTCGTGACCCTGGGCGCTGCGGCAGGGGGCGTGCTGGGCGGCT 294

      67 heLeuIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeu 83
     295 GGCTCCTGGACCGTGCAGGGCGCAAGCTGAGCCTCCTGCTCTGCACCGTG 344

      84 ValLeuLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTr 100
     345 CCCTTCGTGACCGGTTTTGCTGTCATCACCGCGGCCCGGGATGTGTGGAT 394

     100 pLeuValLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerM 117
     395 GCTGCTCGGAGGCCGCCTCCTCACCGGCCTAGCCTGCGGAGTCGCCTCAT 444

     117 etAlaCysCysIleTyrValSerGluLeuValGlyProArgGlnArgGly 133
     445 TAGTGGCACCGGTCTATATCTCGGAAATCGCCTACCCAGCCGTTCGAGGA 494

     134 ValLeuValSerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSe 150
     495 CTGCTCGGCTCCTGTGTGCAGCTGATGGTTGTCACTGGCATCCTCTTGGC 544

     150 rTyrAlaLeuAsnTyrAlaLeuAlaGlyThrProTrpGlyTrpArgHisM 167
     545 CTACGTGGCAGGCTGGGTCCTA.....GAATGGCGCTGGTTGGCTG 585

     167 etPheGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSerLeuLeu 183
     586 TGCTGGGC......TGTGTGCCCCCCACCCTCATGCTGCTGCTCATGTGC 629

     184 PheLeuProAla..... 187
     630 TACATGCCCGAGACCCCACGCTTCCTCCTGACGCAACACCAGTACCAGGA 679

     188 .....GlyThrAspGluThrAlaThrH 195
     680 GGCCATGGCTGCCCTGCGCTTCCTGTGGGGCTCTGAGGAG..... 719

     195 isLysAspLeuIleProLeuGlnGlyGlyGluAlaProLysLeuGlyPro 211
     720 .....GGCTGGGAAGAGCCCCCTGTTGGGGCT 746

     212 GlyArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAspAsnMe 228
     747 GAGCACCAGGGCTTCCAGCTGGCTATGCTG......AGGCGCCCTGGTGT 790

     228 tArgGlyArgThrThrValGlyLeuGlyLeuValLeuPheGlnGlnLeuT 245
     791 CCACAAGCCCCTCATCATCGGCATTTGCCTCATGGTCTTCCAGCAGCTGT 840

     245 hrGlyGlnProAsnValLeuCysTyrAlaSerThrIlePheSerServal 261
     841 CAGGGGTCAACGCCATCATGTTCTATGCCAACACCATCTTTGAGGAGGCC 890

     262 GlyPheHisGlyGlySerSerAlaValLeuAlaSerValGlyLeuGlyAl 278
```

```
 891 AAGTTCAAGGACAGCAGC.........CTGGCCTCGGTCACTGTGGGCAT 931

 278 aValLysValAlaAlaThrLeuThrAlaMetGlyLeuValAspArgAlaG 295
     :::::::|||       |||   :::|||       ::::::||||||||||
 932 CATCCAGGTTCTGTTCACTGCTGTGGCGGCCCTCATCATGGACAGAGCAG 981

 295 lyArgArgAlaLeuLeuLeuAlaGlyCysAlaLeuMetAlaLeuSerVal 311
     |||||:::   ||||||       :::   :::::::|||:::::|||::: 
 982 GGCGAAAACTGCTTCTGGCCTTGTCGGGTGTGATCATGGTTTTCAGCATG 1031

 312 SerGlyIleGlyLeuValSerPheAlaValProMetAspSerGlyProSe 328
     |||:::::::|||       ::::::             :::||||||||||||
1032 AGCGCCTTTGGT......ACCTACTTCAAACTGACCCAGAGTGGCCCCAG 1075

 328 rCysLeuAlaValProAsnAlaThrGlyGlnThrGlyLeuProGlyAspS 345
     |               |||::::::
1076 C..............AACTCCTCC.....................CATG 1089

 345 erGlyLeuLeuGlnAspSerSerLeuProProIleProArgThrAsnGlu 361
       |||||||||                          :::|||   :::   |||
1090 TAGGCCTCCTG.....................GTGCCCATCTCCGCAGAG 1118

 362 AspGlnArgGluProIleLeuSerThrAlaLysLysThrLysProHisPr 378

1118 ................................................. 1118

 378 oArgSerGlyAspProSerAlaProProArgLeuAlaLeuSerSerAlaL 395

1118 ................................................. 1118

 395 euProGlyProProLeuProAlaArgGlyHisAlaLeuLeuArgTrpThr 411
                     ||||||      |||      |||   |||
1119 ................CCTGCTGATGTTCACCTGGGGCTGGCCTGGCTG 1151

 412 AlaLeuLeuCysLeuMetValPheValSerAlaPheSerPheGlyPheGl 428
     |||:::      :::   :::|||:::::::::|||:::   |||:::||
1152 GCTGTAGGCAGCATGTGCCTCTTCATCGCTGGTTTTGCAGTAGGCTGGGG 1201

 428 yProValThrTrpLeuValLeuSerGluIleTyrProValGluIleArgG 445
     ||||:::   ||||||:::::::|||||||||:::|||:::::::|||:::|
1202 ACCCATCCCCTGGCTCCTCATGTCAGAGATCTTCCCTCTGCACATCAAGG 1251

 445 lyArgAlaPheAlaPheCysAsnSerPheAsnTrpAlaAlaAsnLeuPhe 461
     ||   |||   :::   |||      -  ||||||         ::::::
1252 GTGTGGCTACCGGCGTCTGTGTCCTCACCAACTGGTTCATGGCCTTTCTG 1301

 462 IleSerLeuSerPhe.........LeuAspLeuIleGlyThrIleGlyLe 475
     ::::::   :::|||          ::::::::::::          |||
1302 GTGACCAAAGAGTTTAACAGCATCATGGAGATCCTCAGACCCTACGGCGC 1351

 475 uSerTrpThrPheLeuLeuTyrGlyLeuThrAlaValLeuGlyLeuGlyP 492
       |||             :::   :::   :::||||||          |
1352 CTTCTGGCTCACCGCTGCCTTCTGTATCCTCAGCGTCCTT.........T 1392

 492 heIleTyrLeuPheValProGluThrLysGlyGlnSerLeuAlaGluIle 508
     ||         ||||||||||||||||||||:::::::|||  .:::|||
1393 TCACGCTCACCTTTGTCCCTGAGACTAAAGGCAGGACTCTGGAACAAATC 1442

 509 AspGlnGlnPheGlnLysArgArgPheThrLeuSerPheGlyHisArgGl 525
         :::|||:::  |||      |||||||||
1443 ACAGCCCATTTCGAGGGACGGTGACGGACCCTTTCTGTG........... 1481

 525 nAsnSerThrGlyIlePro 531
       ||||||   |||
1482 .......ACTGGCAGCCCT 1493

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC37492

seq_documentation_block:
ID   AAC37492 standard; DNA; 1837 BP.
XX
AC   AAC37492;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 17579.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
```

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
Quality: 482.50 Length: 592
Ratio: 1.546 Gaps: 17
Percent Similarity: 52.703 Percent Identity: 26.182

alignment_block:
US-09-652-292-2 x AAC37492 ..

Align seg 1/1 to: AAC37492 from: 1 to: 1837

```
      4 SerProProValLeuProLeuCysAlaSerValSerLeuLeuGlyGlyLe 20
        :::|||   ||||||      |||:::   |||:::   :::||||||||
    188 ACTCCCTTTGTTCTCTTCACTTGCGTT...GTTGCTGCCATGGGCGGTCT 234

     20 uThrPheGlyTyrGluLeuAlaValIleSerGlyAlaLeu...LeuPro. 35
        |   |||||||||:::::::::::::   :::|||:::       :::|||
    235 CATCTTCGGATACGATATCGGAATCTCCGGTGGTGTGACGTCGATGCCGT 284

     36 .....LeuGlnLeuAspPheGlyLeuSerCysLeuGluGlnGlu..... 48
             |||:::      |||               :::|||:::
    285 CTTTCCTCAAGCGATTCTTCCCGTCGGTGTATCGGAAACAACAAGAGGAC 334
```

```
 49 .....PheLe 50
335 GCGTCAACGAACCAGTACTGTCAGTACGATAGCCCGACGCTAACGATGTT 384

 50 uValGlySerLeuLeuLeuGlyAlaLeuLeuAlaSerLeuValGlyGlyP 67
385 CACATCGTCTCTATATCTAGCGGCGCTAATTTCGTCGCTGGTGGCTTCCA 434

 67 heLeuIleAspCysTyrGlyArgLysGlnAlaIleLeuGlySerAsnLeu 83
435 CCGTGACAAGAAAGTTCGGACGGCGGCTCTCGATGCTCTTCGGCGGCATA 484

 84 ValLeuLeuAlaGlySerLeuThrLeuGlyLeuAlaGlySerLeuAlaTr 100
485 CTCTTCTGCGCCGGAGCTCTCATTAATGGTTTCGCCAAACATGTTTGGAT 534

100 pLeuValLeuGlyArgAlaValValGlyPheAlaIleSerLeuSerSerM 117
535 GCTCATCGTCGGTCGTATCTTGCTTGGTTTCGGTATCGGTTTCGCTAATC 584

117 etAlaCysCysIleTyrValSerGluLeuValGlyProArgGlnArgGly 133
585 AGGCTGTGCCACTGTACCTCTCTGAGATGGCTCCATACAAATACAGAGGA 634

134 ValLeuValSerLeuTyrGluAlaGlyIleThrValGlyIleLeuLeuSe 150
635 GCTTTAAACATTGGTTTCCAACTCTCAATTACAATCGGAATCCTCGTCGC 684

150 rTyrAlaLeuAsnTyrAlaLeuAlaGlyThrPro......TrpGlyTrpA 165
685 CGAAGTGCTCAACTACTTCTTCGCCAAGATCAAAGGCGGTTGGGGATGGC 734

165 rgHisMetPheGlyTrpAlaThrAlaProAlaValLeuGlnSerLeuSer 181
735 GGCTCAGTCTCGGAGGCGCGGTGGTTCCTGCCTTGATCATAACCATCGGC 784

182 LeuLeuPheLeuPro..... 186
785 TCCCTCGTCCTCCCTGACACTCCCAATTCAATGATCGAGCGTGGCCAACA 834

187 .....AlaGlyThrAspGluThrA 193
835 CGAAGAAGCCAAAACCAAGCTCAGACGAATCCGTGGTGTCGATGACGTCA 884

193 laThrHisLysAspLeuIleProLeuGlnGlyGlyGluAlaProLysLeu 209
885 GCCAAGAGTTTGACGATTTGGTCGCCGCTAGTAAAGAGTCGCAGTCGATA 934

210 GlyProGlyArgProArgTyrSerPheLeuAspLeuPheArgAlaArgAs 226
935 .....GAGCACCCGTGGAGAAACCTCCTCCGCCGCAAA.. 967

226 pAsnMetArgGlyArgThrThrValGlyLeuGlyLeuValLeuPheGlnG 243
968 ....TACCGACCACATCTCACTATGGCCGTTATGATTCCGTTCTTTCAAC 1013

243 lnLeuThrGlyGlnProAsnValLeuCysTyrAlaSerThrIlePheSer 259
1014 AGCTAACCGGAATCAATGTGATTATGTTTTACGCTCCGGTTTTGTTCAAC 1063

260 SerValGlyPheHisGlyGlySerSerAlaValLeuAlaSerValGlyLe 276
1064 ACCATTGGTTTCACGACCGATGCTTCT...CTCATGTCCGCTGTGGTCAC 1110

276 uGlyAlaValLysValAlaAlaThrLeuThrAlaMetGlyLeuValAspA 293
1111 TGGCTCGGTTAACGTGGCCGCTACGCTTGTTTCTATCTACGGTGTTGACA 1160

293 rgAlaGlyArgArgAlaLeuLeuLeuAlaGlyCysAlaLeuMetAlaLeu 309
1161 GATGGGGACGTCGGTTTCTCTTTCTTGAAGGTGGTACACAAATGCTTATA 1210

310 SerValSerGlyIleGlyLeuValSerPheAlaValProMetAspSerGl 326

1210 ..... 1210

 326 yProSerCysLeuAlaValProAsnAlaThr.....GlyG 338
1211 .....TGCCAGGCTGTGGTTGCAGCTTGCATAGGGGCCAAGTTTGGGG 1253

 338 lnThrGlyLeuProGlyAspSerGlyLeuLeuGlnAspSerSerLeuPro 354
1254 TAGACGGGACCCCTGGTGAG..... 1273

 355 ProIleProArgThrAsnGluAspGlnArgGluProIleLeuSerThrAl 371
1273 ..... 1273

 371 aLysLysThrLysProHisProArgSerGlyAspProSerAlaProProA 388
1273 ..... 1273

 388 rgLeuAlaLeuSerSerAlaLeuProGlyProProLeuProAlaArgGly 404
1273 ..... 1273

 405 HisAlaLeuLeuArgTrpThrAlaLeuLeuCys.....LeuMetVa 418
1274 .....CTACCGAAGTGGTATGCTATAGTGGTTGTAACGTTCATTTGCAT 1317

 418 lPheValSerAlaPheSerPheGlyPheGlyProValThrTrpLeuValL 435
1318 CTATGTGGCGGGTTTTGCGTGGTCGTGGGGCCCACTAGGGTGGTTAGTAC 1367

 435 euSerGluIleTyrProValGluIleArgGlyArgAlaPheAlaPheCys 451
1368 CGAGTGAAATCTTCCCGTTGGAGATAAGGTCGGCGGCGCAGAGTATCACC 1417

 452 AsnSerPheAsnTrpAlaAlaAsnLeuPheIleSerLeuSerPheLeuAs 468
1418 GTGTCCGTGAACATGATCTTCACGTTCATTATCGCGCAAATCTTCTTGAC 1467

 468 pLeuIleGlyThrIle.....GlyLeuSerTrpThrPheLeuLeuTyrG 483
1468 GATGCTTTGTCATTTGAAGTTTGGGTTA.....TTCCTCGTTTTCG 1508

 483 lyLeuThrAlaValLeuGlyLeuGlyPheIleTyrLeuPheValProGlu 499
1509 CCTTTTTCGTGGTGGTGATGTCGATCTTTGAATACATTTTCTTGCCGGAG 1558

 500 ThrLysGlyGlnSerLeuAlaGluIleAspGlnGlnPheGlnLysArg.. 515
1559 ACGAAAGGGATTCCGATAGAGGAGATGGGTCAAGTGTGGAGGTCACACTG 1608

 516 .....ArgPhe.....ThrLeuSerP 521
1609 GTATTGGTCAAGGTTTGTGGAGGATGGTGAGTATGGGAATGCGCTTGAGA 1658

 521 heGlyHisArgGlnAsnSerThrGly 529
1659 TGGGCAAGAACAGTAACCAAGCTGGA 1684
```

```
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAD09552

seq_documentation_block:
ID   AAD09552 standard; cDNA; 2080 BP.
XX
AC   AAD09552;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human transporter and ion channel-1 (TRICH-1) cDNA.
XX
KW   Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
KW   gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
KW   hypertension; angina; neurological disorder; asthma; bipolar disorder;
```

```
KW   dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;
KW   Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
KW   Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
KW   demyelinating disease; mental disorder; Schizophrenia; polymyositis;
KW   muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
KW   dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
KW   rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
KW   sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
KW   scleroderma; pulmonary artery stenosis; nootropic; Addison's disease;
KW   malabsorption syndrome; hypercholesterolaemia; cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             46..1479
FT                   /*tag= a
FT                   /product= "Human TRICH-1 protein"
XX
PN   WO200146258-A2.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-US35095.
XX
PR   23-DEC-1999;   99US-0172000.
PR   14-JAN-2000; 2000US-0176083.
PR   21-JAN-2000; 2000US-0177332.
PR   28-JAN-2000; 2000US-0178572.
PR   02-FEB-2000; 2000US-0179758.
PR   10-FEB-2000; 2000US-0181625.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Baughn MR,  Burford N,  Au-Young J,  Lu DAM,  Yang J,  Reddy R,  Lal P;
PI   Hillman JL,  Azimzai Y,  Yue H,  Nguyen DB,  Yao MG,  Gandhi AR;
PI   Tang YT,  Khan FA;
XX
DR   WPI; 2001-418042/44.
DR   P-PSDB; AAE04888.
XX
PT   Novel human transporter and ion channel proteins useful for treating
PT   and preventing transport, neurological, muscle and immunological
PT   disorders -
XX
PS   Claim 5; Page 140-141; 160pp; English.
XX
CC   The present sequence is transporter and ion channel-1 (TRICH;1) cDNA.
CC   TRICH is used as vaccine. TRICH is useful for treating a disease or
CC   condition associated with decreased expression of functional TRICH,
CC   such as transport disorder including amyotrophic lateral sclerosis,
CC   cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth
CC   disease, Duchenne muscular dystrophy, angina and hypertension,
CC   neurological disorders including Alzheimer's disease, amnesia, bipolar
CC   disorder, dementia, depression, epilepsy, ischaemic cerebrovascular
CC   disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
CC   disease and Parkinson's disease, demyelinating diseases, mental disorders
CC   including mood, anxiety, Schizophrenia and seasonal affective disorder,
CC   muscle disorder including cardiomyopathy, myocarditis, polymyositis,
CC   dermatomyositis, arrhythmias and asthma and immunological disorders
CC   including AIDS, adult respiratory distress syndrome (ARDS), allergies,
CC   anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
CC   syndrome, systemic lupus erythematosus and other diseases including
CC   sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
CC   artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
CC   glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
CC   psoriasis and viral, bacterial, fungal, helminthic and protozoal
CC   infections. TRICH DNA is useful in gene therapy and in diagnostic
CC   purposes.
XX
SQ   Sequence 2080 BP; 300 A; 688 C; 632 G; 460 T; 0 other;

alignment_scores:
```

```
            Quality:  478.00              Length:    536
              Ratio:  1.654                 Gaps:     15
  Percent Similarity:  53.918   Percent Identity: 27.799

alignment_block:
 US-09-652-292-2 x AAD09552   ..

 Align seq 1/1  to: AAD09552  from: 1  to: 2080

     10 LeuCysAlaSerValSerLeuLeuGlyGlyLeuThrPheGlyTyrGluLe 26
        |||:::|||   ::::::   ||||||   |||:::|||||||:::   ||
    127 CTCGCCGCCTTCGCCGCTGCCCTGGGCCCACTCAGCTTCGGCTTCGCGCT 176

     26 uAlaValIleSerGlyAlaLeuLeuProLeuGln..............  37
        |:::      |||   |||:::      ||||||
    177 CGGCTACAGCTCCCCGGCCATCCCTAGCCTGCAGCGCGCCGCGCCCCCGG 226

     38 ........LeuAsp.................PheGlyLeuSerCysLeu 45
                ||||||                 ||||||
    227 CCCCGCGCCTGGACGACGCCGCCGCCTCCTGGTTCGGG............ 264

     46 GluGlnGluPheLeuValGlySerLeuLeuLeuGlyAlaLeuLeuAlaSe 62
                        :::   :::   ||||||||      :::::
    265 .................GCTGTCGTGACCCTGGGTGCCGCGGCGGGGGG 296

     62 rLeuValGlyGlyPheLeuIleAspCysTyrGlyArgLysGlnAlaIleL 79
        :::::::|||||:::|||:::|||      ||||||||   ::::::|
    297 AGTGCTGGGCGGCTGGCTGGTGGACCGCGCCGGGCGCAAGCTGAGCCTCT 346

     79 euGlySerAsnLeuValLeuLeuAlaGlySerLeuThrLeuGlyLeuAla 95
        ||      ::::::   ::::::||||||      ::::::      |||
    347 TGCTGTGCTCCGTGCCCTTCGTGGCCGGCTTTGCCGTCATCACCGCGGCC 396

     96 GlySerLeuAlaTrpLeuValLeuGlyArgAlaValValGlyPheAlaIl 112
           ::::::     |||:::   ||||||   ::::::|||:::|||
    397 CAGGACGTGTGGATGCTGCTGGGGGGCCGCCTCCTCACCGGCCTGGCCTG 446

    112 eSerLeuSerSerMetAlaCysCysIleTyrValSerGluLeuValGlyP 129
         :::::::::|||:::::::::   :::|||:::||||||::::::   |
    447 CGGTGTTGCCTCCCTAGTGGCCCCGGTCTACATCTCCGAAATCGCCTACC 496

    129 roArgGlnArgGlyValLeuValSerLeuTyrGluAlaGlyIleThrVal 145
        ||      ||||||:::|||   |||      :::      :::::|||
    497 CAGCAGTCCGGGGGTTGCTCGGCTCCTGTGTGCAGCTAATGGTCGTCGTC 546

    146 GlyIleLeuLeuSerTyrAlaLeuAsnTyrAlaLeuAlaGlyThrProTr 162
        ||||||||||||:::|||      :::::::::|||           ||
    547 GGCATCCTCCTGGCCTACCTGGCAGGCTGGGTGCTG.........GAGTG 587

    162 pGlyTrpArgHisMetPheGlyTrpAlaThrAlaProAlaValLeuGlnS 179
        |   |||      ::::::|||          :::|||      |||:::
    588 GCGCTGGCTGGCTGTGCTGGGC......TGCGTGCCCCCCTCCCTCATGC 631

    179 erLeuSerLeuLeuPheLeuProAlaGlyThrAspGluThr......... 192
          |||   :::   |||:::|||            ||||||
    632 TGCTTCTCATGTGCTTCATGCCC...........GAGACCCCGCGCTTC 669

    193 ......AlaThrHisLysAspLeuIleProLeu................. 201
              :::   |||:::          :::
    670 CTGCTGACTCAGCACAGGCGCCAGGAGGCCATGGCCGCCCTGCGGTTCCT 719

    202 .............GlnGlyGlyGluAlaProLysLeuGlyProGlyArgP 214
                     ||||||   |||   |||   :::|||
    720 GTGGGGCTCCGAGCAGGGCTGGGAAGACCCCCCCATCGGG.........G 760

    214 roArgTyrSerPhe...LeuAspLeuPheArgAlaArgAspAsnMetArg 229
           :::   ||||||   |||   |||:::|||   :::   ::::::
    761 CTGAGCAGAGCTTTCACCTGGCCCTGCTGCGG...CAGCCCGGCATCTAC 807

    230 GlyArgThrThrValGlyLeuGlyLeuValLeuPheGlnGlnLeuThrGl 246
                     :::|||::::::|||:::   ||||||||||||:::||
    808 AAGCCCTTCATCATCGGCGTCTCCCTGATGGCCTTCCAGCAGCTGTCGGG 857
```

```
 246 yGlnProAsnValLeuCysTyrAlaSerThrIlePheSerSerValGlyP 263
 858 GGTCAACGCCGTCATGTTCTATGCAGAGACCATCTTTGAAGAGGCCAAGT 907

 263 heHisGlyGlySerSerAlaValLeuAlaSerValGlyLeuGlyAlaVal 279
 908 TCAAGGACAGCAGC.....CTGGCCTCGGTCGTCGTGGGTGTCATC 948

 280 LysValAlaAlaThrLeuThrAlaMetGlyLeuValAspArgAlaGlyAr 296
 949 CAGGTGCTGTTCACAGCTGTGGCGGCTCTCATCATGGACAGAGCAGGGCG 998

 296 gArgAlaLeuLeuLeuAlaGlyCysAlaLeuMetAlaLeuSerValSerG 313
 999 GAGGCTGCTCCTGGTCTTGTCAGGTGTGGTCATGGTGTTCAGCACGAGTG 1048

 313 lyIleGlyLeuValSerPheAlaValProMetAspSerGlyProSerCys 329
1049 CCTTCGGC..... 1056

 330 LeuAlaValProAsnAlaThrGlyGlnThrGlyLeuProGlyAspSerGl 346

1056 ..... 1056

 346 yLeuLeuGlnAspSerSerLeuProProIleProArgThrAsnGluAspG 363

1056 ..... 1056

 363 lnArgGluProIleLeuSerThrAlaLysLysThrLysProHisProArg 379
1057 .....GCCTACTTCAAGCTGACCCAG..... 1077

 380 SerGlyAspProSerAlaProProArgLeuAlaLeuSerSerAlaLeuPr 396
1078 ...GGTGGCCCTGGCAACTCCTCGCACGTGGCCATCTCGGCGCCTGTCTC 1124

 396 oGlyProProLeuProAlaArgGlyHisAlaLeuLeuArgTrpThrAlaL 413
1125 TGCACAGCCTGTTGATGCC......AGCGTGGGGCTGGCCTGGCTGGCCG 1168

 413 euLeuCysLeuMetValPheValSerAlaPheSerPheGlyPheGlyPro 429
1169 TGGGCAGCATGTGCCTCTTCATCGCCGGCTTTGCGGTGGGCTGGGGGCCC 1218

 430 ValThrTrpLeuValLeuSerGluIleTyrProValGluIleArgGlyAr 446
1219 ATCCCCTGGCTCCTCATGTCAGAGATCTTCCCTCTGCATGTCAAGGGCGT 1268

 446 gAlaPheAlaPheCysAsnSerPheAsnTrpAlaAlaAsnLeuPheIleS 463
1269 GGCGACAGGCATCTGCGTCCTCACCAACTGGCTCATGGCCTTTCTCGTGA 1318

 463 erLeuSerPheLeuAspLeuIleGlyThrIleGlyLeuSerTrpThrPhe 479
1319 CCAAGGAGTTCAGCAGCCTCATGGAGGTCCTCAGGCCCTATGGAGCCTTC 1368

 480 LeuLeuTyrGlyLeuThrAlaValLeuGlyLeuGlyPheIleTyrLeuPh 496
1369 TGGCTTGCCTCCGCTTTCTGCATCTTCAGTGTCCTTTTCACTTTGTTCTG 1418

 496 eValProGluThrLysGlyGlnSerLeuAlaGluIleAspGlnGlnPheG 513
1419 TGTCCCTGAAACTAAAGGAAAGACTCTGGAACAAATCACAGCCCATTTTG 1468

 513 lnLysArg 515
1469 AGGGGCGA 1476
```